



1 Source attribution using FLEXPART and carbon monoxide 2 emission inventories: SOFT-IO version 1.0

3 Bastien Sauvage¹, Alain Fontaine¹, Sabine Eckhardt³, Antoine Auby⁴, Damien Boulanger²,
4 Hervé Petetin¹, Ronan Paugam⁵, Gilles Athier¹, Jean-Marc Cousin¹, Sabine Darras³, Philippe
5 Nédélec¹, Andreas Stohl³, Solène Turquety⁶, Jean-Pierre Cammas⁷ and Valérie Thouret¹.

6
7 ¹Laboratoire d'Aérodynamique, Université de Toulouse, CNRS, UPS, France

8 ²Observatoire Midi-Pyrénées, Toulouse, France

9 ³NILU - Norwegian Institute for Air Research, Kjeller, Norway

10 ⁴CAP HPI, Leeds, United Kingdom

11 ⁵King's College, London, United Kingdom

12 ⁶Laboratoire de Météorologie Dynamique/IPSL, UPMC Univ. Paris 6, Paris, France

13 ⁷Observatoire des Sciences de l'Univers de la Réunion (UMS 3365) et Laboratoire de l'Atmosphère et des
14 Cyclones (UMR 8105), Université de la Réunion, Saint-Denis, La Réunion, France

15

16

17 *Correspondence to:* Bastien Sauvage (bastien.sauvage@aero.obs-mip.fr)

18 **Abstract.** Since 1994, the In-service Aircraft for a Global Observing System (IAGOS) program has produced
19 in-situ measurements of the atmospheric composition during more than 51000 commercial flights. In order to
20 help analyzing these observations and understanding the processes driving the observed concentration
21 distribution and variability, we developed the SOFT-IO tool to quantify source/receptor links for all measured
22 data. Based on the FLEXPART particle dispersion model (Stohl et al., 2005), SOFT-IO simulates the
23 contributions of anthropogenic and biomass burning emissions from the ECCAD emission inventory database
24 for all locations and times corresponding to the measured carbon monoxide mixing ratios along each IAGOS
25 flight. Contributions are simulated from emissions occurring during the last 20 days before an observation,
26 separating individual contributions from the different source regions. The main goal is to supply added-value
27 products to the IAGOS database by evincing the geographical origin and emission sources driving the CO
28 enhancements observed in the troposphere and lower stratosphere. This requires a good match between observed
29 and modeled CO enhancements. Indeed, SOFT-IO detects more than 95% of the observed CO anomalies over
30 most of the regions sampled by IAGOS in the troposphere. In the majority of cases, SOFT-IO simulates CO
31 pollution plumes with biases lower than 10-15 ppbv. Differences between the model and observations are larger
32 for very low or very high observed CO values. The added-value products will help in the understanding of the
33 trace-gas distribution and seasonal variability. They are available in the IAGOS data base via
34 <http://www.iagos.org>. The SOFT-IO tool could also be applied to similar data sets of CO observations (e.g.
35 ground-based measurements, satellite observations). SOFT-IO could also be used for statistical validation as well
36 as for inter-comparisons of emission inventories using large amounts of data.

37 1 Introduction

38 Tropospheric pollution is a global problem caused mainly by natural or human-triggered biomass burning,
39 and anthropogenic emissions related to fossil fuel extraction and burning. Pollution plumes can be transported



40 quickly on a hemispheric scale (within at least 15 days) by large scale winds or, more slowly (Jacob, 1999),
41 between the two hemispheres (requiring more than 3 months). Global anthropogenic emissions are for some
42 species (CO₂) in constant increase (Boden et al., 2015). However, recent commitments of some countries to
43 reduce greenhouse gas emissions (e.g. over the U.S., U.S. EPA's Inventory of U.S. Greenhouse Gas Emissions
44 and Sinks, 1990-2013; <http://www.epa.gov/climatechange/ghgemissions/usinventoryreport.html>) seems to
45 induce a stalling in other global emissions (NO_x, SO₂ and Black Carbon, Stohl et al., 2015), except for some
46 regions (Brazil, Middle East India, China) where NO_x emissions increase (Miyazaki, 2017). In order to better
47 understand large-scale pollution transport, large amounts of in situ and space-based data have been collected in
48 the last three decades, allowing a better understanding of pollution variability and its connection with
49 atmospheric transport patterns (e.g. Liu et al., 2013). These data-sets are also useful to quantify global pollution
50 evolution with respect to the emissions trends described above.

51 Despite the availability of large trace gas data sets, the data interpretation remains difficult for the following
52 reasons: (1) the sampling mode does not correspond to an a priori defined scientific strategy, as opposed to data
53 collected during field campaigns; (2) the statistical analysis of the data can be complicated by the large number
54 of different sources contributing to the measured pollution, and an automated analysis of the contributions from
55 these different sources is required if, for instance, regional trends in emissions are to be investigated; (3) the
56 sheer size of some of the data sets can make the analysis rather challenging. Among the long-term pollution
57 measurement programs, the IAGOS airborne program (<http://www.iagos.org/>, formerly known as the
58 Measurement of OZone by Airbus In-service airCRAFT -MOZAIC- program) is the only one delivering in-situ
59 measurement data from the free troposphere. IAGOS provides regular global measurements of ozone (O₃) - since
60 1994 -, carbon monoxide (CO) - since 2002 -, and nitrogen oxides (NO_y) - for the period 2001-2005 - obtained
61 during more than 51000 commercial aircraft flights up to now, with substantial extent of the instrumented
62 aircraft recently. The analysis of the IAGOS database is also complicated by the fact that primary pollutants (CO
63 and part of NO_y) are emitted by multiple sources, while secondary compounds (O₃) are produced by
64 photochemical transformations of these pollutants, often most efficiently when pollutants from different sources
65 mix.

66 A common approach to separate the different sources influencing trace gas observations is based on the
67 determination of the air mass origins through Lagrangian modeling. This approach allows linking the emission
68 sources to the trace gas observations (e.g. Nédélec et al., 2005; Sauvage et al., 2005, 2006; Tressol et al. 2008;
69 Gressent et al. 2014; Clark et al., 2015; Yamasoe et al., 2015). Lagrangian modeling of the dispersion of
70 particles allows accounting efficiently for processes such as large-scale transport, turbulence and convection.
71 When coupled with emission inventories Lagrangian modeling of passive tracers allows for instance to
72 understand ozone anomalies (Cooper et al., 2006; Wen et al., 2012), to quantify the importance of lightning NO_x
73 emissions for tropospheric NO₂ columns measured from space (Beirle et al., 2006), to investigate the origins of
74 O₃ and CO over China (Ding et al., 2013), or to investigate the sources influencing the observed CO₂ over the
75 high northern latitudes (Vay et al., 2011).

76 To help analyzing a large data set such as the IAGOS observations, it is important to provide scientific users
77 a tool for characterizing air mass transport and emission sources. This study presents a methodology to
78 systematically establish a link between emissions sources (biomass burning and anthropogenic emissions) and
79 concentrations at the receptor locations. Since CO is a substance that is emitted by combustion sources (both



80 anthropogenic and biomass burning) and since CO has a lifetime of months in the troposphere (Logan et al.,
81 1981; Mauzerall et al., 1998), it is often used as a tracer for pollution transport (Staudt et al. 2001; Yashiro et al.,
82 2009; Barret et al., 2016). It is therefore convenient to follow past examples and use simulated CO source
83 contributions to gauge the influence of pollution sources on the measurements also with SOFT-IO. Our
84 methodology uses the FLEXPART Lagrangian particle dispersion model (Stohl et al., 2005) and emission
85 inventories from the ECCAD emission database (Granier et al., 2012) in order to quantify the influence of
86 emissions sources on the IAGOS CO measurements. The goal is to provide the scientific community with added
87 value products that will help them analyzing and interpreting the large number of IAGOS measurements. The
88 methodology has the benefit to be adaptable to multiple emission inventories without re-running FLEXPART
89 simulations. It is also easily adaptable to analyze other datasets of trace gas measurements such as from ground
90 based observations, sondes, aircraft campaigns or satellite observations.

91 The methodology will be described in the next section, and then evaluated at the example of case studies of
92 pollution plumes observed by IAGOS aircraft. Further evaluation is performed through statistical analysis.
93 Finally we discuss the limitations of the methodology by estimating its sensitivity to different input data sets
94 (emission inventories, meteorological analyses).

95 **2. In-situ observations database: MOZAIC and IAGOS programs**

96 The MOZAIC program (Marengo et al., 1998) was initiated in 1993 by European scientists, aircraft
97 manufacturers and airlines to better understand the natural variability of the chemical composition of the
98 atmosphere and how it is changing under the influence of human activity, with particular interest in the impact of
99 aircraft exhaust. Between August 1994 and November 2014, MOZAIC performed airborne in-situ measurements
100 of ozone, water vapor, carbon monoxide, and total nitrogen oxides. The measurements are geolocated (latitude,
101 longitude and pressure) and come along with meteorological observations (wind direction and speed,
102 temperature). Data acquisition is performed automatically during round-trip international flights (ascent, descent
103 and cruise phases) from Europe to America, Africa, Middle East, and Asia (Fig. 1).

104 Based on the technical expertise of MOZAIC, the IAGOS program (Petzold et al., 2015, and references therein)
105 has taken over and provides observations since July 2011. The IAGOS data set still includes ozone, water vapor,
106 carbon monoxide, meteorological observations, and measurements of cloud droplets (number and size) are also
107 performed. Depending on optional additional instrumentation, measurements of nitrogen oxides, total nitrogen
108 oxides or, in the near-future, greenhouse gases (CO₂ and CH₄), or aerosols, will also be made.

109 Since 1994, the IAGOS-MOZAIC observations have created a big data set that is stored in a single database
110 holding data from more than 51000 flights. The data set can be used by the entire scientific community, allowing
111 studies of chemical and physical processes in the atmosphere, or validation of global chemistry transport models
112 and satellite retrievals. Most of the measurements have been collected in the upper troposphere and lower
113 stratosphere, between 9 and 12 km altitude, with 500 flights/ aircraft/ year on up to 7 aircraft up to now.

114
115 The MOZAIC and IAGOS data (called “IAGOS” from here on) used in this study are in-situ observations of CO
116 only, which is being measured regularly on every aircraft since 2002 with more than 30000 flights, using a
117 modified infrared filter correlation monitor (Nédélec et al., 2003; Nédélec et al., 2015). The accuracy of the CO
118 measurements has been estimated at (30 s response time) ± 5 ppb, or $\pm 5\%$.



119

120 Several case studies of CO pollution plumes (Table 1) using IAGOS data have been published, where model
121 simulations allowed attribution of the measured CO enhancements to anthropogenic or biomass burning
122 emissions, either measured in the boundary layer or in the free troposphere, following regional or synoptic-scale
123 transport (e.g. Nédélec et al., 2005; Tressol et al., 2008; Cammas et al., 2009; Elguindi et al., 2010). These case
124 studies are used here to better define the requirements for our methodology (meteorological analyses and
125 emission inventory inputs). Some of them are detailed and re-analyzed in Sect. 4.

126 3. Estimation of carbon monoxide source regions: methodology

127 To establish systematic source-receptor relationships for IAGOS observations of CO, the Lagrangian dispersion
128 model FLEXPART (Stohl et al., 1998, 2005; Stohl and Thomson, 1999) is run over the entire database.
129 Lagrangian dispersion models usually represent the differential advection better than global Eulerian models
130 (which do not well resolve intercontinental pollution transport; Eastham et al., 2017), at a significantly lower
131 computational cost. In particular, small-scale structures in the atmospheric composition can often be
132 reconstructed from large-scale global meteorological data, which makes model results comparable to high-
133 resolution in situ observations (Pisso et al., 2010). In the past, many studies (Nédélec et al., 2005 ; Tressol et
134 al., 2008; Cammas et al., 2009; Elguindi et al., 2010; Gressent et al., 2014) used FLEXPART to investigate
135 specific pollution events observed by the IAGOS aircraft. However, in these former case studies, the link
136 between sources and observations of pollution was guessed a priori. The transport model was then used to
137 validate the hypothesis. For example, in the Cammas et al. (2009) study, observations of high CO during summer
138 in the upper troposphere and lower stratosphere east of Canada were guessed to originate from biomass burning
139 over Canada as this region is often associated with pyro-convection whose intensity usually peaks in the
140 summer. This origin was confirmed by the model analysis. In general, the origin of the observed pollution cannot
141 be guessed a priori, especially when analyzing measurements from thousands of flights. Moreover, multiple
142 sources are most of the time involved when the observed pollution is the result of the mixing of polluted air
143 masses from different regions and source types.

144 CO is often used as a tracer to quantify the contributions of the different sources to the observed pollution
145 episodes. CO is emitted by both the combustion of fossil fuels and by biomass burning, and its photochemical
146 lifetime against OH attack is usually 1 to 2 months in the troposphere (Logan et al., 1981; Mauzerall et al.,
147 1998). Therefore it is possible to link elevated CO mixing ratios (with respect to its seasonally varying
148 hemispheric baseline) to pollution sources without simulating the atmospheric chemistry.

149 3.1 Backward transport modeling

150 Simulations were performed using the version 9 of FLEXPART, which is described in detail by Stohl et al.
151 (2005) (and references therein). The model was driven using wind fields from the European Centre for Medium-
152 Range Weather Forecast (ECMWF) 6-hourly operational analyses and 3-hour forecasts. The ECMWF data are
153 gridded with a $1^\circ \times 1^\circ$ horizontal resolution, and with a number of vertical levels increasing from 60 in 2002 to
154 137 since 2013. The model was also tested using higher horizontal resolution (0.5°), and with ECMWF ERA-
155 Interim reanalysis, as their horizontal and vertical resolution and model physics are homogeneous during the
156 whole period of IAGOS CO measurements. However, operational analyses were used for our standard set-up, as



157 the transport model reproduced CO better when using these data for several case studies of pollution transport,
158 especially for plumes located in the UT. Indeed, operational analyses provide a better vertical resolution since
159 2006 (91 levels until 2013, then 137 levels against 60 levels for ERA-Interim) and thus a better representation of
160 the vertical wind shear, and the underlying meteorological model is also more modern than the one used for
161 producing ERA-Interim. Vertical resolution is obviously the most critical factor for modeling such CO plumes
162 with the best precision in terms of location and intensity (Eastham and Jacob, 2017).

163 Using higher horizontal resolution for met-fields analyses and forecasts (0.5° vs 1°) showed no influence on the
164 simulated carbon monoxide, despite larger computational time and storage needs. We assume further
165 improvement can be obtained using even higher horizontal resolution (0.1°), but this was not feasible at this
166 stage and should be considered in the future.

167

168 In order to be able to represent the small-scale structures created by the wind shear and observed in many
169 IAGOS vertical profiles, the model is initialized along IAGOS flight tracks every 10 hPa during ascents and
170 descents, and every 0.5° in latitude and longitude at cruise altitude. This procedure leads to *i* model initialization
171 boxes along every flight track. For each *i*, 1000 particles are released. Indeed 1000 to 6000 particles are
172 suggested for correct simulations in similar studies based on sensitivity tests on particles number (Wen et al.,
173 2012; Ding et al., 2013). For instance, a Frankfurt (Germany) to Windhoek (Namibia) flight contains around 290
174 boxes (290000 particles) of initialization as a whole.

175 FLEXPART is set up for backward simulations (Seibert and Frank, 2004) from these boxes as described in Stohl
176 et al. (2003) and backward transport is computed for 20 days prior to the in-situ observation, which is sufficient
177 to consider hemispheric scale pollution transport in the mid-latitudes (Damoah et al., 2004; Stohl et al., 2002;
178 Cristofanelli et al., 2013). This duration is also expected to be longer than the usual lifetime of polluted plumes
179 in the free troposphere, i.e. the time when the concentration of pollutants in plumes is significantly larger than
180 the surrounding background. Indeed, the tropospheric mixing time scale has been estimated to be typically
181 shorter than 10 days (Good et al., 2003; Pissu et al., 2009). Therefore the model is expected to be able to link air
182 mass anomalies such as strong enhancements in CO to the source regions of emissions (Stohl et al., 2003). It is
183 important to note that we aim to simulate recent events of pollution explaining CO enhancements over the
184 background, but not to simulate the CO background which results from aged and well-mixed emissions.

185 The FLEXPART output is a residence time, as presented and discussed in Stohl et al. (2003). These data
186 represent the average time spent by the transported air masses in a grid cell, divided by the air density, and are
187 proportional to the sensitivity of the receptor mixing ratio to surface emissions. In our case, it is calculated for
188 every input point along the flight track, every day for $N_t = 20$ days backward in time, on a 1° longitude x 1°
189 latitude global grid with $N_z = 12$ vertical levels (every 1 km from 0 to 12 km, and 1 layer above 12 km).

190 Furthermore, the altitude of the 2 PVU potential vorticity level above or below the flight track is extracted from
191 the wind and temperature fields, in order to locate the CO observations above or below the dynamical tropopause
192 according to the approach of Thouret et al. (2006).

193 3.2 Emission inventories from the ECCAD project

194 The main goal of the Emissions of atmospheric Compounds & Compilation of Ancillary Data (ECCAD) project
195 (Granier et al., 2012) is to provide scientific and policy users with datasets of surface emissions of atmospheric



196 compounds and ancillary data, i.e. data required for estimating or quantifying surface emissions. All the emission
 197 inventories and ancillary data provided by ECCAD are published in the scientific literature.

198 For the current study, we selected five CO emission inventories. Four of them are available at global scale
 199 (MACCity and EDGAR v4.2 for anthropogenic; GFED 4 and GFAS v1.2 -GFAS v1.0 for 2002- for fires) from
 200 the ECCAD database and cover most of the IAGOS CO database presented here (2002 - 2013). The global scale
 201 inventories have a $0.1^\circ \times 0.1^\circ$ to $0.5^\circ \times 0.5^\circ$ horizontal resolution. They are provided with daily, monthly or
 202 yearly time resolution. They are listed in Table 2 along with the references describing them. The four global
 203 inventories are used to study the model's performance and sensitivity in Sect. 5.

204 To further test the sensitivity to the emission inventories, we also used one regional inventory, which is expected
 205 to provide a better representation of emissions in its region of interest than generic global inventories. For
 206 biomass burning, the International Consortium for Atmospheric Research on Transport and Transformation
 207 (ICARTT) campaign's North American emissions inventory developed by Turquety et al. (2007) for the summer
 208 of 2004 and provided at $1^\circ \times 1^\circ$ horizontal resolution was tested. It combines daily area burned data from forest
 209 services with the satellite data used by global inventories, and uses a specific vegetation database, including
 210 burning of peat lands which represent a significant contribution to the total emissions.

211 3.3 Coupling transport output with CO emissions

212 Calculating the recent contributions $C(i)$ (kg m^{-3}) of CO emissions for every one of the i model's initialization
 213 points along the flight tracks requires three kinds of data:

- 214 • the residence time T_R (in seconds, gridded with $N_x = 360$ by $N_y = 180$ horizontal points, $N_z = 12$ vertical
 215 levels, $N_t = 20$ days) from backward transport described in Sect. 3.1,
- 216 • CO surface emissions $E_{CO}(N_x, N_y, N_t)$ (in $\text{kg CO} / \text{m}^2 / \text{s}$)
- 217 • the injection profile $Inj(z)$ defining the fraction of pollutants diluted in the different vertical levels (with
 218 Δz being the thickness, in meters) just after emissions:

$$220 \text{ (Eq. 1) } C(i) = \sum_{t=1}^{N_t} \sum_{y=1}^{N_y} \sum_{x=1}^{N_x} \sum_{z=1}^{N_z} Inj(z) \frac{T_R(x, y, z, t, i) E_{CO}(x, y, t)}{\Delta z(z)}$$

221

222 In the case of anthropogenic emissions, CO is simply emitted into the first vertical layer of the residence time
 223 grid ($\Delta z = 1000\text{m}$).

224

225 For biomass burning emissions, in the tropics and mid latitudes regions, the lifting of biomass burning plumes is
 226 usually due to small and large scale dynamical processes, such as turbulence in the boundary layer, deep
 227 convection and frontal systems, which are usually represented by global meteorological models. At higher
 228 latitudes, however, boreal fires can also be associated with pyro-convection and quick injection above the
 229 planetary boundary layer. Pyro-convection plume dynamics are often associated with small-scale processes that
 230 are not represented in global meteorological data and emission inventories (Paugam et al 2016). In order to
 231 characterize the effect of these processes, we implemented three methodologies to parameterize biomass
 232 injection height:



- 233 • the first one (named DENTENER) depends only on the latitude and uses constant homogeneous
234 injection profiles as defined by Dentener et al. (2006), i.e. 0-1 km for the tropics [30S-30N] (see green
235 line in Fig. 2), 0-2 km for the mid-latitudes [60S-30S, 30N-60N] (see blue line in Fig. 2) and 0-6 km for
236 the boreal regions [90S-60S, 60N-90N] (not shown in Fig. 2).
- 237 • the second named MIXED uses the same injection profiles as in DENTENER for the tropics and mid-
238 latitudes, but for the boreal forest, injection profiles are deduced from a lookup table computed with the
239 plume rise model PRMv2 presented in Paugam et al. (2015). Using PRMv2 runs for all fires from
240 different years of the Northern-American MODIS archive, three daily Fire Radiative Power (FRP)
241 classes (under 10 TJ/day, between 10 and 100 TJ/day, and over 100 TJ/day) were used to identify three
242 distinct injection height profiles (see brown, red, and black lines in Fig. 2). Although PRMv2 reflects
243 both effects of the fire intensity through the input of FRP and active fire size and effects of the local
244 atmospheric profile, here for sake of simplicity only FRP is used to classify the injection profile.
245 Furthermore, when applied to the IAGOS data set, the MIXED method uses equivalent daily FRP
246 estimated from the emitted CO fluxes given by the emission inventories as described in Kaiser et al.
247 (2012)
- 248 • the third method named hereafter APT uses homogeneous profile defined by the daily plume top
249 altitude as estimated for each 0.1x0.1 pixel of the GFAS v1.2 inventory available for 2003 to 2013
250 (Rémy et al. 2016, and http://www.gmes-atmosphere.eu/oper_info/global_nrt_data_access/gfas_ftp/).
251 As in the MIXED method, GFAS v1.2 is using the plume model PRMV2 from Paugam et al. (2015),
252 but here the model is run globally for every assimilated GFAS-FRP pixel.
253

254 3.4 Automatic detection of CO anomalies

255 For individual measurement cases, plumes of pollution can most of the time be identified by the human eye
256 using the observed CO mixing ratio time series or the CO vertical profiles. However, this is not feasible for a
257 database of tens of thousands of observation flights. In order to create statistics of the model's performance, we
258 need to systematically identify observed pollution plumes in the IAGOS database. The methodology to do this is
259 based on what has been previously done for the detection of layers in the MOZAIC database (Newell et al.,
260 1999; Thouret et al., 2000), along with more recent calculations of the CO background and CO percentiles define
261 for different regions along the IAGOS data set (Gressent et al., 2014). An example demonstrating the procedure,
262 which is described below, is shown in Fig. 3.

263

264 In a first step, the measurement time series along the flight track (number of measurements n_{TOT}) is separated
265 into three parts:

- 266 1. Ascent and descent vertical profiles (n_{VP}) in the PBL (altitudes ranging from the ground to 2 km) and in
267 the free troposphere (from 2 km to the top altitude of the vertical profiles),
- 268 2. measurements at cruising altitude in the upper troposphere (n_{UT}),
- 269 3. measurements in the lower stratosphere (n_{LS})

270 such that
$$n_{TOT} = n_{VP} + n_{UT} + n_{LS}$$



271 where n_{VP} , n_{UT} and n_{LS} are the number of measurements along tropospheric ascents and descents, and in the upper
 272 troposphere and lower stratosphere, respectively. A range of altitudes from the surface to a top altitude identifies
 273 vertical profiles. The top altitude is 75 hPa above the 2 pvu dynamical tropopause (Thouret et al., 2006) when
 274 the aircraft reaches/leaves cruising altitude (during ascent/descent). The PV is taken from the ECMWF
 275 operational analyses and evaluated at the aircraft position by FLEXPART. Observations made during the cruise
 276 phase are flagged as upper tropospheric if the aircraft is below the 2 pvu dynamical tropopause. If not,
 277 observations are considered as stratospheric and then are ignored in the rest of the paper. Although CO
 278 contributions are calculated also in the stratosphere, the present study focuses on tropospheric pollution only.

279

280 In a second step, the CO background mixing ratio is determined for each tropospheric part (C_{VP_back} and C_{UT_back} ,
 281 see Fig. 3 for illustration) for the tropospheric vertical profiles and for the upper troposphere respectively. For
 282 tropospheric vertical profiles, the linear regression of CO mixing ratio versus altitude is calculated from 2 km to
 283 the top of the vertical profiles, to account for the usual decrease of background CO with altitude. Data below
 284 2 km are not used because high CO mixing ratios caused by fresh emissions are usually observed close to surface
 285 over continents. The slope a (in ppb m^{-1}) of the linear regression is used to determine the background so that
 286 $C_{VP_back} = aZ$. The background is removed from the C_{VP} tropospheric vertical profiles mixing ratio to obtain a
 287 residual CO mixing ratio C_{VP}^R (Eq. 2).

$$288 \quad (\text{Eq. 2): } C_{VP}^R = C_{VP} - C_{VP_back},$$

289

290 For the upper troposphere, the CO background mixing ratio (C_{UT_back}) is determined using seasonal median
 291 values (over the entire IAGOS database) for the different regions of Figure 4. Note that this approach was not
 292 feasible for vertical profiles as for most of the visited airports there are not enough data to establish seasonal
 293 vertical profiles. As for the profiles, background values are subtracted from the UT data to obtain residual C_{UT}^R
 294 (Eq. 3):

$$295 \quad (\text{Eq. 3): } C_{UT}^R = C_{UT} - C_{UT_back}$$

296

297 In a third step, CO anomalies C^A are determined for tropospheric vertical profiles (C_{VP}^A) and in the upper
 298 troposphere (C_{UT}^A). Residual C_{VP}^R and C_{UT}^R values are flagged as CO anomalies when these values exceed the
 299 third quartile (Q3) of the residual mixing ratio $C_{VP}^R(Q3)$ for vertical profiles, or the third quartile of the residual
 300 seasonal values $C_{UT_season}^R(Q3)$ in the different regions (Fig. 4) for the UT. Note that $C_{VP}^R(Q3)$ or $C_{UT_season}^R(Q3)$
 301 needs to be higher than 5 ppb (the accuracy of the CO instrument; Nédélec et al., 2015) in order to consider an
 302 anomaly:

$$303 \quad (\text{Eq. 4): } C_{VP}^A = C_{VP}^R \text{ if } C_{VP}^R > C_{VP}^R(Q3)$$

$$304 \quad (\text{Eq. 5): } C_{UT}^A = C_{UT}^R \text{ if } C_{UT}^R > C_{UT_season}^R(Q3)$$

305 In the examples shown in Fig. 3a and Fig. 3b, the red line represents CO anomalies.

306 With this algorithm CO plumes are automatically detected in the entire IAGOS database. For each identified
 307 plume, minimum and maximum values of the date, latitude, longitude and altitude, as well as the CO mean and
 308 maximum mixing ratio, are archived. These values are used for comparison with modeled CO values.

309



310 4. Selected case studies to evaluate CO emission inventories and SOFT-IO's performance

311 As described in Sect. 2, a number of case studies documented in the literature were selected from the IAGOS
312 database in order to get a first impression of the model's performance. These case studies have been chosen to
313 represent the different pollution situations that are often encountered in the troposphere in terms of emissions
314 (anthropogenic or biomass burning) and transport (at regional or synoptic scale, pyro-convection, deep
315 convection, frontal systems). Systematic evaluation of the model performance against emission inventories will
316 be presented in Sect. 5.

317 4.1 Anthropogenic emission inventories

318 Among the case studies listed in Table 1, four were selected in order to illustrate the evaluation of the inventories
319 used for anthropogenic emissions:

- 320 • Landing profiles over Hong Kong from 19th of July and 22nd of October 2005 were selected in order to
321 investigate specifically Asian anthropogenic emissions.
- 322 • During the 10th of March 2002 Frankfurt–Denver and 27th of November 2002 Dallas–Frankfurt flights,
323 IAGOS instruments observed enhanced CO plumes in the North Atlantic upper troposphere, also linked
324 to anthropogenic emissions.

325 Figure 5a shows the observed (black line) and simulated (colored lines) CO mixing ratios above Hong Kong
326 during 22nd of October 2005. Note that background is not simulated but estimated from the observations as
327 described in Sect.3.4 (blue line, C_{VP_back}). The dashed blue line represents the residual CO mixing ratio C_{VP}^R .
328 Observations show little variability in the free troposphere down to around 3 km. Strong pollution is observed
329 below, with + 300 ppb enhancement over the background on average between 0 and 3 km. Note that we do not
330 discuss CO enhancement above 3 km.

331 In agreement with C_{VP}^R , SOFT-IO simulates a strong CO enhancement in the lowest 3 km of the profile, caused
332 by fresh emissions. However, the simulated enhancement is less strong than the observed one, a feature that is
333 typical for this region, as we shall see later.

334 In addition to the CO mixing ratio, SOFT-IO calculates CO source contributions and geographical origins of the
335 modeled CO, respectively displayed in Fig. 5b and Fig. 5c (using the methodology described in Sec. 3.4) and
336 using here MACCity and GFAS v1.2 as example. For the geographical origin we use the same 14 regions as
337 defined for the GFED emissions (<http://www.globalfiredata.org/data.html>). Note that only the average of the
338 calculated CO is displayed for each anomaly (0-3km; 3.5-6km) in Fig. 5b and Fig. 5c.

339

340 Colored lines in Fig. 5a show the calculated CO using anthropogenic sources described by the two inventories
341 selected in Sect. 3.2, MACCity (green line) and EDGARv4.2 (yellow line), along the flight track. In both cases,
342 biomass burning emissions are described by GFASv1.2. Emissions from fires have negligible influence (less
343 than 3%) on this pollution event as depicted in Fig. 5b.

344 In the two simulations, the calculated CO mixing ratio is below 50 ppb in the free troposphere, as we do not
345 simulate background concentrations with SOFT-IO. CO enhancement around 4 to 6 km is overestimated by
346 SOFT-IO. CO above 6 km is not considered as an anomaly, as $C_{UT}^R < C_{UT_season}^R(Q3)$. Simulated mixing ratios in
347 the 0-2 km polluted layer are almost homogeneous, with values around 280 ppb using MACCity and around 160
348 ppb using EDGARv4.2. They are attributed to anthropogenic emissions (more than 97% of the simulated CO)



349 originating mostly from Central Asia with around 95% influence. In this regard, the CO simulated using
350 MACCity is in better agreement with the observed CO than the one obtained using EDGARv4.2. Indeed, using
351 MACCity, simulated CO reaches 90% of the observed enhancement (+ 300 ppb on average) over the background
352 (around 100 ppb), while for EDGARv4.2 the corresponding value is only 53%, indicating strong underestimation
353 of this event. The difference in the calculated CO using these two inventories is also consistent with the results
354 of Granier et al. (2011) who showed strong discrepancies in the Asian anthropogenic emissions in different
355 inventories.

356

357 Figure 6a shows the CO measurements at cruising altitude during a transatlantic flight between Frankfurt and
358 Denver on 10th of March 2002. The dashed blue line represents the residual CO C_{UT}^R . Observations indicate that
359 the aircraft encountered several polluted air masses with CO mixing ratios above 110 to 120 ppb, which are the
360 seasonal median CO values in the two regions visited by the aircraft, obtained from the IAGOS database (see
361 Gressent et al., 2014). Three pollution plumes are measured:

- 362 • around 100°W (around +10 ppb of CO enhancement on average): plume 1
- 363 • between 80°W and 50°W (+30 ppb of CO enhancement on average): plume 2
- 364 • between 0° and 10°E (+40 ppb of CO enhancement on average): plume 3.

365 These polluted air masses are surrounded by stratospheric air masses with CO values lower than 80-90 ppb. As
366 polluted air masses were sampled at an altitude of around 10 km, they are expected to be due to long-range
367 transport of pollutants.

368 The calculated CO is shown in Fig. 6a using MACCity (green line), EDGARv4.2 (yellow line) for anthropogenic
369 emissions and GFASv1.0 for biomass burning emissions. SOFT-IO estimates that these plumes are mostly
370 anthropogenic (representing 77% to 93% of the total simulated CO, Fig. 6b). Pollution mostly originates from
371 Central and South-East Asia, with strong contribution from North America (Fig. 6c) for plume 3.

372 SOFT-IO correctly locates the three observed polluted air masses with the two anthropogenic inventories. CO is
373 also correctly calculated using MACCity, with almost the same mixing ratios on average as the observed
374 enhancements in the three plumes. Only 2/3 of the observed enhancements are simulated using EDGARv4.2,
375 except for plume 1 with better results. We have already seen in the previous case study that emissions in Asia
376 may be underestimated, especially in the EDGARv4.2 inventory.

377 Similar comparisons were performed in the four case studies selected to estimate and validate the anthropogenic
378 emission inventories coupled with the FLEXPART model. Results are summarized in Table 3. For three of the
379 cases, SOFT-IO simulations showed a better agreement with observations when using MACCity than when
380 using EDGARv4.2. In the fourth case both inventories performed equally well. One reason for the better
381 performance of MACCity is the fact that it provides monthly information (Table 2).

382

383 4.2 Biomass burning emission inventories

384 In order to evaluate and choose biomass burning emission inventories, we have selected eleven case studies with
385 fire-induced plumes (Table 1). Seven of them focused on North-American biomass burning plumes observed in
386 the free troposphere above Europe (flights on 30th of June, 22nd and 23rd of July 2004) and in the upper
387 troposphere/lower stratosphere above the North Atlantic (29th of June 2004) (e.g. Elguindi et al., 2010; Cammas



388 et al., 2009). Two are related to the fires over Western Europe during the 2003 heat wave (Tressol et al. 2008).
389 The two last ones, on the 30th and 31st of July 2008, focused on biomass burning plumes observed in the ITCZ
390 region above Africa as described in a previous study (Sauvage et al., 2007a).
391 The three datasets selected to represent biomass burning emissions are based on different approaches: GFAS
392 v1.2 (Kaiser et al., 2012) and GFED 4 emissions (Giglio et al., 2013) are calculated daily. GFAS v1.2 presents
393 higher spatial resolution. The ICARTT campaign inventory (Turquety et al., 2007) was specifically designed for
394 North-American fires during the summer of 2004 with additional input from local forest services.
395 Figure 7a illustrates the calculated CO contributions for the different fire emission inventories for one of the case
396 studies, on 22nd of July 2004 above Paris. The observations (black line) show high levels of CO in an air mass in
397 the free troposphere between 3 and 6 km, with mixing ratios 140 ppb above the background (blue line) deduced
398 from measurements. This pollution was attributed to long-range transport of biomass burning emission in North
399 America by Elguindi et al. (2010). Outside of the plume, the CO concentration decreases with altitude, from
400 around 150 ppb near the ground, to 100 ppb background in the upper free troposphere. This last value
401 corresponds to the median CO seasonal value deduced from the IAGOS database (Gressent et al., 2014). CO is
402 not considered as an anomaly near the ground as $C_{UT}^R < C_{UT,season}^R(Q3)$.
403 SOFT-IO simulations were performed for this case using MACCity to represent anthropogenic emissions, and
404 GFAS v1.2 (green line), GFED 4 (yellow line), or the ICARTT campaign inventory (red line). Fire vertical
405 injection is realized using the MIXED approach for the three biomass burning inventories, in order to only
406 evaluate the impact of choosing different emission inventories. In the three simulations, contributions show two
407 peaks, one near the ground that is half due to local anthropogenic emissions and half due to contributions from
408 North American biomass burning and thus not considered in this discussion.
409 The second more intense peak, simulated in the free troposphere where the enhanced CO air masses were
410 sampled, is mostly caused by biomass burning emissions (87% of the total calculated CO, Fig. 7b), originating
411 from North-America (99% of the total enhanced CO). When calculated using the ICARTT campaign inventory,
412 the simulated CO enhancement reaches over 150 ppb, which is 10 ppb higher than the observed mixing ratio
413 above the background (+140 ppb), but only for the upper part of the plume.
414 When using global inventories, the simulated contribution peak reaches 70 ppb using GFASv1.2 and 100 ppb
415 using GFED4, which appears to underestimate the measured enhancement (+140 ppb) by up to 50% to 70%
416 respectively. This comparison demonstrates the large uncertainty in simulated CO caused by the emission
417 inventories, both in the case of biomass burning or anthropogenic emissions. For that reason we aim to provide
418 simulations with different global and regional inventories in for the IAGOS data set.
419 As the ICARTT campaign inventory was created using local observations in addition to satellite products, the
420 large difference in the simulated CO compared to the other inventories may in part be due to different
421 quantification of the total area burned (for GFED, GFAS using the FRP as constraint). Turquety et al. (2007)
422 also discussed the importance of peat land burning during that summer. They estimated that they contributed
423 more than a third of total CO emissions (11 Tg of the 30 Tg emitted during summer 2004).
424
425 Figure 8a shows CO mixing ratios as a function of latitude for a flight from Windhoek (Namibia) to Frankfurt
426 (Germany) in July 2008. Observations indicate that the aircraft flew through polluted air masses around the
427 equator (10°S to 10°N), with +100 (+125) ppb of CO on average (at the most) above the 90 ppb background



428 deduced from seasonal IAGOS mixing ratios over this region. Such CO enhancements have been attributed to
429 regional fires injected through ITCZ convection (Sauvage et al., 2007b).

430 The SOFT-IO simulations (colored lines in Fig. 8a) link these air masses mostly to recent biomass burning
431 (responsible for 68% of the total simulated CO, Fig. 8b) in South Africa (Fig. 8c). The calculated CO shows
432 similar features both with GFED4 (yellow line) and GFASv1.2 (green line). The simulation also captures well
433 the intensity variations of the different peaks: maximum values around the equator, lower ones south and north
434 of the equator. The most intense simulated CO enhancement around the equator fits the observed CO
435 enhancement of +125 ppb better when using GFED4 (90 ppb) than when using GFASv1.2 (75 ppb). However
436 the comparison also reveals an underestimation of the CO anomaly's amplitude by around 10 ppb to 25 ppb on
437 average by SOFT-IO. The model is thus only able to reproduce 75% to 90% of the peak concentrations on
438 average. Stroppiana et al. (2010) indeed showed that there are strong uncertainties in the fire emission
439 inventories over Africa (164 to 367 Tg CO per year).

440 **5 Statistical evaluation of the modeled CO enhancements in pollution plumes**

441 In this section, we present a statistical validation of the SOFT-IO calculations based on the entire IAGOS CO
442 data base (2003-2013). The ability of SOFT-IO in simulating CO anomalies is evaluated compared to in situ
443 measurements in terms of:

- 444 • spatial and temporal frequency of the plumes
- 445 • mixing ratio enhancements in the plumes

446 To achieve this, SOFT-IO performances are investigated over different periods of IAGOS measurements
447 depending on the emission inventory used. Three of the four global inventories selected previously (MACCcity,
448 GFAS v1.2, GFED4) are available between 2003 and 2013. EDGAR v4.2 ends in 2008. In the following
449 sections (Sect.5.1 and 5.2), we discuss in detail the results obtained with MACCcity and GFAS v1.2 between
450 2003 and 2013. Other emission inventory combinations are discussed in Sect. 5.3 when investigating SOFT-IO
451 sensitivity to input parameters.

452 **5.1 Detection frequency of the observed plumes with SOFT-IO**

453 The ability of SOFT-IO to reproduce CO enhancements was investigated using CO plumes obtained applying the
454 methodology described in Sect. 3.4 on all flights of the IAGOS database between 2003 and 2013. The frequency
455 of simulated plumes that coincide with the observed C^A anomalies is then calculated. Simulated plumes are
456 considered when matching in time and space the observed plumes, while modeled CO is on average higher than
457 5 ppb within the plume. Note that at this stage, we do not consider the intensity of the plumes.

458 The resulting detection rates are presented in Fig. 9 for eight of the eleven regions shown in Fig. 4. Statistics are
459 presented separately for three altitude levels (Lower Troposphere 0-2 km, Middle Troposphere 2-8 km and
460 Upper Troposphere > 8 km). Figure 9 shows that SOFT-IO performance in detecting plumes is very good and
461 not strongly altitude or region-dependent. In the three layers (LT, MT and UT), detection rates are higher than
462 95% and even close to 100% in the LT where CO anomalies are often related to short-range transport. Detection
463 frequency slightly decreases in the MT and the UT where CO modeling accuracy suffers from larger errors in
464 vertical and horizontal transport. On the contrary CO anomalies in the LT are most of the times related to short-
465 range transport of local pollution, which are well represented in SOFT-IO. For four regions we found less good



466 results: South America MT and UT, Africa MT and North Asia UT but with still high detection frequency (82%
467 to 85%). Note that only relatively few plumes (313 to 3761) were sampled by the IAGOS aircraft fleet in these
468 regions.

469

470 **5.2 Intensity of the simulated plumes**

471 The second objective of SOFT-IO is to accurately simulate the intensity of the observed CO anomalies. Fig. 10a
472 displays the bias between the means of the observed and modeled plumes for the regions sampled by IAGOS and
473 in the three vertical layers (LT, MT and UT). As explained above this bias is calculated for the 2003-2013 period
474 and using both anthropogenic emission from MACCity and biomass burning emissions from GFAS v1.2 and the
475 plume detection methodology described in Sect. 3.4.

476 The most documented regions presenting CO polluted plumes (Europe, North America, Africa, North Atlantic
477 UT, Central Asia MT and UT, South America, South Asia UT) present low biases (lower than ± 5 ppb, and up to
478 ± 10 ppb for Central Asia MT, South America UT), which demonstrate a high skill of SOFT-IO.

479 Over several other regions with less frequent IAGOS flights, however, biases are higher, around ± 10 -15 ppb for
480 Africa UT and South Asia MT; around ± 25 -50 ppb for Central Asia LT, South Asia LT and North Asia UT.
481 Except for the last region, the highest biases are found in the Asian lower troposphere, suggesting
482 misrepresentation of local emissions. Indeed there is a rapid increase of emissions in this large area (Tanimoto et
483 al., 2009) associated with high discrepancies between different emission inventories (Wang et al., 2013; Stein et
484 al., 2014) and underestimated emissions (Zhang et al., 2015).

485 It is important to note that the biases remain of the same order (± 10 -15 ppb) when comparing the first (Q1),
486 second (Q2) and third (Q3) quartiles of the CO anomalies observed and modeled within most of the regions (Fig.
487 10b). This confirms the good capacity of the SOFT-IO software in reproducing the CO mixing ratios anomaly in
488 most of the observed pollution plumes.

489 Differences become much larger when considering outlier values of CO anomalies (lower and upper whiskers, \pm
490 2.7σ or 99.3%, Fig. 10b), which means for exceptional events of very low and very high CO enhancements
491 (accounting for 1.4% of the CO plumes), with biases from ± 10 ppb to ± 50 ppb for most of the regions. Higher
492 discrepancies are found in the lower and the upper troposphere and can reach ± 50 to ± 200 ppb in two specific
493 regions (North Asia UT and South Asia LT) for these extreme CO anomalies. Note that North Asia UT and
494 South Asia LT present respectively extreme pollution events related to pyro-convection (Nédélec et al., 2005) for
495 the first region, and to strong anthropogenic surface emissions (Zhang et al., 2012) for the second one. It may
496 suggest that the model fails to correctly reproduce the transport for some specific but rare events of pyro-
497 convection.

498 When looking at the origin of the different CO anomalies (Fig. 10c), most of them are dominated by
499 anthropogenic emissions, which account for more than 70% of the contributions on average, except for South
500 America and Africa, which are strongly influenced by biomass burning (Sauvage et al. 2005, 2007c; Yamasoe et
501 al., 2014). Discussing origins of the CO anomalies in detail is out of the scope of this study, but gives here some
502 information on the model performance. It is interesting to note that two of the three regions most influenced by
503 anthropogenic emissions, South Asia LT and Central Asia LT, with more than 90% of the enhanced CO coming
504 from anthropogenic emissions, are the highest biased regions compared to observations. This is not the case for



505 Europe LT for example, which also has a high anthropogenic influence. As stated before, anthropogenic
506 emissions in Asia are more uncertain than elsewhere (Stein et al., 2014).

507

508 In order to go a step further in the evaluation of SOFT-IO in reproducing CO anomalies mixing ratios, Fig. 11
509 displays the monthly mean time series of the observed (black line) and calculated (blue line) CO anomalies in
510 three vertical layers (LT, MT and UT). This graph provides higher temporal resolution of the anomalies. CO
511 polluted plumes are displayed here using MACCity and GFAS v1.2 over the 2003-2013 periods and for the two
512 regions with the largest number of observed CO anomalies, Europe and North America.

513 It is worth noting the good ability of SOFT-IO in quantitatively reproducing the CO enhancements observed by
514 IAGOS. This is especially noticeable in the LT and UT, with similar CO mixing ratios observed and modeled
515 during the entire period and within the standard deviation. However, the amplitude of the seasonal cycle of CO
516 maxima is highly underestimated (-100%) after January 2009 in the European LT, where anthropogenic sources
517 are predominant with more than 90% influence (Fig. 10c). This suggests misrepresentation of anthropogenic
518 emissions in Europe after the year 2009 (Stein et al., 2014).

519 In the middle troposphere (2-8 km), the CO plumes are systematically overestimated by SOFT-IO by 50% to
520 100% compared to the observations. This might be related to different reasons:

- 521 • the chosen methodology of the CO plume enhancements detection for those altitudes (described in Sect.
522 3.4), which may lead to a large number of plumes with small CO enhancements, which are difficult to
523 simulate. This could be due to the difficulty in defining a realistic CO background in the middle
524 troposphere.
- 525 • the source-receptor transport which may be more difficult to simulate between 2-8 km than in the LT
526 where receptors are close to sources; or than in the UT where most of the plumes are related to
527 convection detrainment better represented in the models than MT detrainment which might be less
528 intense.
- 529 • The frequency of the IAGOS observations which is lower in the MT than in the UT.

530 Correlation coefficients between simulated and observed plumes are highest in the LT (0.56 to 0.79) and lower
531 (0.30 to 0.46) in the MT and in the UT, suggesting some difficulties for the model in lifting up pollution from the
532 surface to the UT.

533 5.3 Sensitivity of SOFT-IO to input parameters

534 Different factors influence the ability of SOFT-IO to correctly reproduce CO pollution plumes. Among them,
535 transport parameterizations (related to convection, turbulence, etc) are not evaluated in this study as they are
536 inherent of the FLEXPART model. In this section, the model sensitivity to the chosen emission inventory is
537 evaluated. For this, a set of sensitivity studies is performed to investigate different configurations of the emission
538 inventories :

- 539 • type of inventory: MACCity, EDGAR for anthropogenic, GFED4, GFAS v1.2 or ICARTT for biomass
540 burning
- 541 • biomass burning injection heights: DENTENER, MIXED or APT approach (detailed in Sect. 3.3).

542



543 SOFT-IO performances are then investigated using Taylor diagrams (Taylor et al. 2001). The methodology
544 (choice of regions, vertical layers, sampling periods) is similar to the one used to analyze the ability of the model
545 to correctly reproduce the frequency and the intensity of the CO plumes with MACCity and GFAS (Sect.5.1 and
546 Sec5.2).

547 **5.3.1 Anthropogenic emission inventories**

548 Sensitivity of SOFT-IO to anthropogenic emissions is investigated between 2002 and 2008, using GFAS with
549 MACCity or EDGARv4.2. Fig. 12a presents a Taylor diagram for the two configurations (dots for MACCity,
550 crosses for EDGAR) for the regions and for the vertical layers described previously (Sect. 5.1 and Sect. 5.2),
551 while Fig. 12b represents the mean bias between each model configuration and the IAGOS observations.

552 As already seen in Sect. 4.1 for the case studies chosen to investigate anthropogenic emissions, slightly better
553 results seem to be obtained with MACCity. The Taylor diagram shows for most of the regions higher
554 correlations and lower biases in this case. These results are not surprising, as MACCity (Granier et al., 2011) is a
555 more recent inventory compared to EDGARv4.2 (Janssens-Maenhout et al., 2010), and expected to better
556 represent anthropogenic emissions. However the differences between the two inventories are most of the time
557 very low, as global emission inventories tend to be quite similar.

558 Regionally, however, results with EDGARv4.2 can be better, such as over South Asia LT and MT, Central Asia
559 MT and North Asia UT. This supports our choice of maintaining several different inventories in SOFT-IO.

560 **5.3.2 Biomass burning emissions**

561 We first investigate the sensitivity of SOFT-IO to the type of biomass burning inventory, using MACCity with
562 GFAS v1.2 or GFED 4 (2003-2013), using the same MIXED methodology for vertical injection of emissions
563 (Fig. 2). As for anthropogenic emissions, Fig. 13 represents the Taylor diagram and averaged biases for the
564 different configurations.

565 Performances (correlations, standard deviations and biases) are very similar for both biomass burning
566 inventories, with smaller differences compared to anthropogenic inventories. Even for regions dominated by
567 biomass burning such as Africa or South America as depicted previously (Fig. 11c), the sensitivity of the SOFT-
568 IO performance to the type of global fire inventory is below 5 ppb.

569

570 Based on case studies, we discussed in Sect. 4.2 the comparison of CO contributions modeled using regional fire
571 emission inventories. It resulted in a better representation of biomass burning plumes using the specifically
572 designed campaign inventory than using the global inventories (Table 4). However, there is no clear evidence of
573 this result when investigating the model performances during the whole summer 2008. On contrary to Sect. 4.2,
574 it is hard to conclude of systematic better results using the ICARTT inventory. While simulations (not shown)
575 give better results for a few specific events of very high CO using ICARTT, similarly good results are obtained
576 when using GFASv1.2 or GFED4 for most other cases. It is worth noting that IAGOS samples biomass burning
577 plumes far from ICARTT sources, after dispersion and diffusion during transport in the atmosphere. Besides,
578 few boreal fire plumes (that would be better represented using ICARTT), are sampled by the IAGOS program.

579



580 Secondly, we investigate the influence of the vertical injection scheme for the biomass burning emissions, using
581 the three methodologies for determining injection heights described in Sect. 3.3. Sensitivity tests (Fig. 13c and
582 Fig 13d) demonstrate a small influence of the injection scheme on the simulated plumes. The largest influence is
583 found over North Asia UT, where pyro-convection has been highlighted in the IAGOS observations (Nédélec et
584 al., 2005), with however less than 5 ppb difference between the different schemes. More generally, small vertical
585 injection influence is probably due to too few cases where boreal fire emissions are injected outside the PBL by
586 pyro-convection, as shown in the Paugam et al. (2016) study, combined with a too low sampling frequency of
587 boreal fire plumes by IAGOS.

588

589 6 Conclusions

590

591 Analyzing long term in situ observations of trace gases can be difficult without a priori knowledge of the
592 processes driving their distribution and seasonal/regional variability, like transport and photochemistry. This is
593 particularly the case for the extensive IAGOS database, which provides a large number of aircraft-based in-situ
594 observations (more than 51000 flights so far) distributed on a global scale, and with no a priori sampling
595 strategy, unlike dedicated field campaigns.

596

597 In order to help studying and analyzing such a large data set of in situ observations, we developed a system that
598 allows quantifying the origin of trace gases both in terms of geographical location as well as source type. The
599 SOFT-IO module (<https://doi.org/10.25326/2>) is based on the FLEXPART particle dispersion model that is run
600 backward from each trace gas observation, and on different emission inventories (EDGAR v4.2, MACCity,
601 GFED 4, GFAS v1.2) than can be easily changed.

602

603 The main advantages of the SOFT-IO module are:

- 604 • Its flexibility. Source-receptor relationships pre-calculated with the FLEXPART particle dispersion
605 model can be coupled easily with different emission inventories, allowing each user to select model
606 results based on a range of different available emission inventories.
- 607 • CO calculation, which is computationally very efficient, can be repeated easily whenever updated
608 emission information becomes available without running again the FLEXPART model. It can also be
609 extended to a larger number of emission datasets, particularly when new inventories become available,
610 or for emission inventories inter-comparisons. It can also be extended to other species with similar or
611 longer lifetime as CO to study other type of pollution sources.
- 612 • High sensitivity of the SOFT-IO CO mixing ratios to source choice for very specific regions and case
613 studies, especially in the LT most of the time driven by local or regional emissions, may also help
614 improving emission inventories estimates through evaluation with a large database such as IAGOS one.
615 Indeed as it is based on a Lagrangian dispersion model, the tool presented here is able to reproduce
616 small-scale variations, which facilitates comparison to in situ observations. It can then be used to
617 validate emission inventories by confronting them to downwind observations of the atmospheric
618 composition, using large database of in situ observations of recent pollution.



- 619 • More generally SOFT-IO can be used in the future for any kind of atmospheric observations (e.g.
620 ground based measurements, satellite instruments, aircraft campaigns) of passive tracers.

621

622 In this study SOFT-IO is applied to all IAGOS CO observations, using ECMWF operational meteorological
623 analysis and 3-hour forecast fields and inventories of anthropogenic and biomass burning emissions available on
624 the ECCAD portal. SOFT-IO outputs are evaluated first at the examples of case studies of anthropogenic and
625 biomass burning pollution events. The evaluation is then extended statistically, for the entire 2003-2013 period,
626 over 14 regions and 3 vertical layers of the troposphere.

627

628 The main results are the following:

- 629 • By calculating the contributions of recent emissions to the CO mixing ratio along the flight tracks,
630 SOFT-IO identifies the source regions responsible for the observed pollution events, and is able to
631 attribute such plumes to anthropogenic and/or biomass burning emissions.
- 632 • On average, SOFT-IO detects 95% of all observed CO plumes. In certain regions, detection frequency
633 reaches almost 100%.
- 634 • SOFT-IO gives a good estimation of the CO mixing ratio enhancements for the majority of the regions
635 and the vertical layers. In majority, the CO contribution is reproduced with a mean bias lower than 10-
636 15 ppb, except for the measurements in the LT of Central and South Asia and in the UT of North Asia
637 where emission inventories seems to be less accurate.
- 638 • CO anomalies calculated by SOFT-IO are very close to observations in the LT and UT where most of
639 the IAGOS data are recorded. Agreement is lower in the MT, possibly because of numerous thinner
640 plumes of lower intensity (maybe linked to the methodology of the plume selection).
- 641 • SOFT-IO has less skill in modeling CO in extreme plume enhancements with biases higher than 50 ppb.

642

643 In its current version, SOFT-IO is limited by different parameters, such as inherent parameterization of the
644 Lagrangian model, but also by input of external parameters such as meteorological field analysis and emission
645 inventories. Sensitivity analyses were then performed using different meteorological analysis and emissions
646 inventories, and are summarized as follow:

- 647 • Model results were not very sensitive to the resolution of the meteorological input data. Increasing the
648 resolution from 1 deg to 0.5 deg resulted only in minor improvements. On the other hand, using
649 operational meteorological analysis allowed more accurate simulations than using ERA-Interim
650 reanalysis data, perhaps related to the better vertical resolution of the former.
- 651 • Concerning anthropogenic emissions sensitivity tests, results display regional differences depending on
652 the emission inventory choice. Slightly better results are obtained using MACCity.
- 653 • Model results were not sensitive to biomass burning global inventories, with good results using either
654 GFED 4 or GFAS v1.2. However, a regional emission inventory shows better results for few individual
655 cases with high CO enhancements. There is a low sensitivity to parameterizing the altitude of fire
656 emission injection, probably because events of fires injected outside of the PBL are rare or because
657 IAGOS does not frequently sample of such events

658



659 Using such CO calculations and partitioning makes it possible to link the trends in the atmospheric composition
660 with changes in the transport pathways and/or changes of the emissions.

661 SOFT-IO products will be made available through the IAGOS central database (<http://iagos.sedoo.fr/#L4Place>)
662 and are part of the ancillary products (<https://doi.org/10.25326/3>)

663

664

665

666

667 **Acknowledgements**

668

669 The authors would like to thank the ECCAD project for providing emission inventories. The authors acknowledge
670 the strong support of the European Commission, Airbus, and the Airlines (Lufthansa, Air-France, Austrian, Air
671 Namibia, Cathay Pacific, Iberia and China Airlines so far) who carry the MOZAIC or IAGOS equipment and
672 perform the maintenance since 1994. In its last 10 years of operation, MOZAIC has been funded by INSU-
673 CNRS (France), Météo-France, Université Paul Sabatier (Toulouse, France) and Research Center Jülich (FZJ,
674 Jülich, Germany). IAGOS has been additionally funded by the EU projects IAGOS-DS and IAGOS-ERI. The
675 MOZAIC-IAGOS database is supported by AERIS (CNES and INSU-CNRS). The former CNES-ETHER
676 program has funded this project.

677

678

679

680 **References**

681

682 Barret, B., Sauvage, B., Bennouna, Y., and Le Flochmoen, E.: Upper-tropospheric CO and O₃ budget during the
683 Asian summer monsoon, *Atmos. Chem. Phys.*, 16, 9129–9147, doi:10.5194/acp-16-9129-2016, 2016

684 Beirle, S.; Spichtinger, N.; Stohl, A.; et al.: Estimating the NO(x) produced by lightning from GOME and NLDN
685 data: a case study in the Gulf of Mexico, *Atm. Chem. Phys.*, 6, 1075–1089, 2006.

686 Boden, T.A., G. Marland, and R.J. Andres. 2015. Global, Regional, and National Fossil-Fuel CO₂ Emissions.
687 Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak
688 Ridge, Tenn., U.S.A. doi 10.3334/CDIAC/00001_V2015, 2015

689 Cammas, J.-P., Brioude, J., Chaboureau, J.-P., Duron, J., Mari, C., Mascart, P., N'ed'elec, P., Smit, H., Pätz,
690 H.W., Volz-Thomas, A., Stohl, A., and Fromm, M.: Injection in the lower stratosphere of biomass fire emissions
691 followed by long-range transport: a MOZAIC case study, *Atm. Chem. Phys.*, 9, 5829–5846, [http://www. atmos-](http://www.atmos-chem-phys.net/9/5829/2009/)
692 [chem-phys.net/9/5829/2009/](http://www.atmos-chem-phys.net/9/5829/2009/), 2009.

693 Clark, Hannah, Bastien Sauvage, Valerie Thouret, Philippe Nedelec, Romain Blot, Kuo-Ying Wang, Herman
694 Smit, et al.: The First Regular Measurements of Ozone, Carbon Monoxide and Water Vapour in the Pacific
695 UTLS by IAGOS, *Tellus B*, 67. doi:10.3402/tellusb.v67.28385, 2015.

696 Cooper, O. R.; Stohl, A.; Trainer, M.; et al : Large upper tropospheric ozone enhancements above midlatitude
697 North America during summer: In situ evidence from the IONS and MOZAIC ozone measurement network, *J.*
698 *Geophys. Res.*, 111, D24, 2006.

699 Cristofanelli, P., Fierli, F., Marinoni, A., Calzolari, F., Duchi, R., Burkhart, J., Stohl, A., Maione, M., Arduini, J.,
700 and Bonasoni, P.: Influence of biomass burning and anthropogenic emissions on ozone, carbon monoxide and
701 black carbon at the Mt. Cimone GAW-WMO global station (Italy, 2165 m a.s.l.), *Atmos. Chem. Phys.*, 13, 15-
702 30, doi:10.5194/acp-13-15-2013, 2013

703 Damoah, R., Spichtinger, N., Forster, C., James, P., Mattis, I., Wandinger, U., Beirle, S., Wagner, T., and Stohl,
704 A.: Around the world in 17 days -hemispheric-scale transport of forest fire smoke from Russia in May 2003,
705 *Atm. Chem. Phys.*, 4, 1311–1321, 2004.

706 Dentener, F., Kinne, S., Bond, T., Boucher, O., Cofala, J., Generoso, S., Ginoux, P., Gong, S., Hoelzemann, J. J.,
707 Ito, A., Marelli, L., Penner, J. E., Putaud, J.-P., Textor, C., Schulz, M., van der Werf, G. R., and Wilson, J.:
708 Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom,
709 *Atmos. Chem. Phys.*, 6, 4321–4344, doi:10.5194/acp-6-4321-2006, 2006

710 Ding, A., T. Wang, and C. Fu (2013), Transport characteristics and origins of carbon monoxide and ozone in
711 Hong Kong, South China, *J. Geophys. Res. Atmos.*, 118, 9475–9488, doi:10.1002/jgrd.50714, 2013

712 Eastham, S. D. and Jacob, D. J.: Limits on the ability of global Eulerian models to resolve intercontinental
713 transport of chemical plumes, *Atmos. Chem. Phys.*, 17, 2543–2553, doi:10.5194/acp-17-2543-2017, 2017.

714 Elguindi, N., Clark, H., Ordonez, C., Thouret, V., Flemming, J., Stein, O., Huijnen, V., Moinat, P., Inness, A.,
715 Peuch, V.-H., Stohl, A., Turquety, S., Athier, G., Cammas, J.-P., and Schultz, M.: Current status of the ability of
716 the GEMS/MACC models to reproduce the tropospheric CO vertical distribution as measured by MOZAIC,
717 *Geosci. Model Dev.*, 3, 501–518, <http://www.geosci-model-dev.net/3/501/2010/>, 2010.

718 Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Silva Dias, M. A. F., Andreae, M. O., Prins, E., Santos, J.
719 C., Gielow, R., and Carvalho Jr., J. A.: Including the sub-grid scale plume rise of vegetation fires in low



- 720 resolution atmospheric transport models, *Atmos. Chem. Phys.*, 7, 3385-3398, doi:10.5194/acp-7-3385-2007,
721 2007
- 722 Giglio, S., Randerson, J.T., Van der Werf, G.R.: Analysis of daily, monthly, and annual burned area using the
723 fourth-generation global fire emissions database (GFED4), *J. Geophys. Res.*, 10.1002/jgrg.20042, 2013
- 724 Good, P., Giannakopoulos, C., O'Connor, F.M., Arnold, S.R., de Reus, M., Schlager, H.: Constraining
725 tropospheric mixing timescales using airborne observations and numerical models, *Atm. Chem. Phys.*, 3, 1023-
726 1035, 2003.
- 727 Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., Denier van der Gon, H., Frost, G., Heil, A., Kaiser, J.,
728 Kinne, S., Klimont, Z., Kloster, S., Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, T.,
729 Raut, J.-C., Riahi, K., Schultz, M., Smith, S., Thompson, A., van Aardenne, J., van der Werf, G., and van
730 Vuuren, D.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional
731 scales during the 1980-2010 period, *Climatic Change*, 109, 163–190, <http://dx.doi.org/10.1007/s10584-011-0154-1>,
732 0154-1, 10.1007/s10584-011-0154-1, 2011.
- 733 Granier, C., Damas, S., Liousse, C., Middleton, P., Mieville, A., et al. : The ECCAD Database: Emissions of
734 Atmospheric Compounds & Compilation of Ancillary Data. *IGAC Newsletter*, pp.18-20, 2012
- 735 Gressent, A., Sauvage, B., Defer, E. et al. : Lightning NO_x influence on large-scale NO_y and O-3 plumes
736 observed over the northern mid-latitudes, *Tellus B*, 66, 25544, 2014
- 737 Hanna, S. R.: Applications in air pollution modeling, *Atmospheric Turbulence and Air Pollution Modelling*,
738 1982.
- 739 Jacob, D.J.: *Introduction to Atmospheric Chemistry*, Princeton University Press, 1999
- 740 Janssens-Maenhout, G., Petrescu, A. M. R., Muntean, M., and Blujdea, V.: Verifying Greenhouse Gas
741 Emissions: Methods to Support International Climate Agreements, *Greenhouse Gas Measurement and*
742 *Management*, 2010.
- 743 Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J. J., Razinger, M.,
744 Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire
745 assimilation system based on observed fire radiative power, *Biogeosciences*, 9, 527–554, 2012.
- 746 Liu, L., Logan, J.A., Murray, L.T., Pumphrey, H.C., Schwartz, M.J., Megretskaya, I.A.: Transport analysis and
747 source attribution of seasonal and interannual variability of CO in the tropical upper troposphere and lower
748 troposphere, *Atm. Chem. Phys.*, 13, 129-146, 2013.
- 749 Logan, J.A., Prather, M.J., Wofsy, S.C. et al. : Tropospheric Chemistry – A Global Perspective, *J. Geophys.*
750 *Res.*, 86, 7210-7254, 1981.
- 751 Marengo, A; Thouret, V; Nedelec, P; et al.: Measurement of ozone and water vapor by Airbus in-service aircraft:
752 The MOZAIC airborne program, An overview, *J. Geophys. Res.*, 103, D19, 1998.
- 753 Mauzerall, DL; Logan, JA; Jacob, DJ; et al. : Photochemistry in biomass burning plumes and implications for
754 tropospheric ozone over the tropical South Atlantic, *J. Geophys. Res.*, 103, D7, 1998.
- 755 Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., and Kanaya, Y.: Decadal changes in global
756 surface NO_x emissions from multi-constituent satellite data assimilation, *Atmos. Chem. Phys.*, 17, 807-837,
757 doi:10.5194/acp-17-807-2017, 2017



- 758 Nédélec, P., Thouret, V., Brioude, J., Sauvage, B., Cammas, J. P., and Stohl, A.: Extreme CO concentrations in
759 the upper troposphere over northeast Asia in June 2003 from the in situ MOZAIC aircraft data, *Geophys. Res.*
760 *Letts.*, 32, 2005.
- 761 Nedelec, P.; Cammas, JP; Thouret, V; et al: An improved infrared carbon monoxide analyser for routine
762 measurements aboard commercial Airbus aircraft: technical validation and first scientific results of the MOZAIC
763 III programme, *Atm. Chem. Phys.*, 3, 1551-1564, 2003
- 764 Nedelec, P., Blot, R., Boulanger, D. et al.: Instrumentation on commercial aircraft for monitoring the
765 atmospheric composition on a global scale: the IAGOS system, technical overview of ozone and carbon
766 monoxide measurements, *Tellus B*, 67, 27791, 2015.
- 767 Newell, R.E., Thouret, V., Cho, J.Y.N., Stoller, P., Marenco, A., and Smit, H.G.S.: Ubiquity of quasi-horizontal
768 layers in the atmosphere, *Nature*, 398, 316-319, doi:10.1038/18642, 1999
- 769 Paugam, R., Wooster, M., Atherton, J., Freitas, S. R., Schultz, M. G., and Kaiser, J. W.: Development and
770 optimization of a wildfire plume rise model based on remote sensing data inputs – Part 2, *Atmos. Chem. Phys.*
771 *Discuss.*, 15, 9815-9895, doi:10.5194/acpd-15-9815-2015, 2015
- 772 Paugam, R., Wooster, M., Freitas, S. R., and Val Martin, M.: A review of approaches to estimate wildfire plume
773 injection height within large scale atmospheric chemical transport models, *Atmos. Chem. Phys.*, 16, 907-925,
774 doi:10.5194/acpd-16-907-2016, 2016
- 775 Petzold, A., Thouret, V., Gerbig, C. et al.: Global-scale atmosphere monitoring by in-service aircraft –current
776 achievements and future prospects of the European Research infrastructure IAGOS, *Tellus B*, 67, 28452, 2015
- 777 Pisso, I., Real, E., Law, K.S., Legras, B., Bousseres, N., Attié, J.L., Schlager, H.: Estimation of mixing in the
778 troposphere from Lagrangian traec gas reconstructions during long-range pollution plume transport , *J. of*
779 *Geophys. Res.* , 114, D19301, 2010.
- 780 Rémy, S., Veira, A., Paugam, R., Sofiev, M., Kaiser, J. W., Marenco, F., Burton, S. P., Benedetti, A., Engelen,
781 R. J., Ferrare, R., and Hair, J. W.: Two global climatologies of daily fire emission injection heights since 2003,
782 *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2015-1048, in review, 2016
- 783 Rodean, H. C.: *Stochastic Lagrangian Models of Turbulent Diffusion*, vol. 26, American Meteorological Society,
784 1996.
- 785 Sauvage, B., Thouret, V., Cammas, J.P., Gheusi, F., Athier, G., and Nédélec P.: Tropospheric ozone over
786 Equatorial Africa: regional aspects from the MOZAIC data, *Atmos. Chem. Phys.*, 5, 311-335, 2005
- 787 Sauvage, B., V. Thouret, A. M. Thompson, J. C. Witte, J.-P. Cammas, P. Nédélec, and G. Athier: Enhanced
788 view of the “tropical Atlantic ozone paradox” and “zonal wave one” from the in situ MOZAIC and SHADOZ
789 data, *J. Geophys. Res.*, 111, D01301, doi:[10.1029/2005JD006241](https://doi.org/10.1029/2005JD006241), 2006.
- 790 Sauvage, B.; Thouret, V.; Cammas, J. -P.; et al.: Meridional ozone gradients in the African upper troposphere,
791 *Geophys. Res. Letts.*, 34, L03817, 2007a
- 792 Sauvage, B., R. V. Martin, A. van Donkelaar, and J. R. Ziemke: Quantification of the factors controlling tropical
793 tropospheric ozone and the South Atlantic maximum, *J. Geophys. Res.*, 112, D11309,
794 doi:10.1029/2006JD008008, 2007b
- 795 Sauvage B., Martin R.V., van Donkelaar A., Liu X., Chance K., Jaeglé L., Palmer P.I., Wu S. , Fu T.-M. :
796 Remote sensed and in situ constraints on processes affecting tropical ozone, *Atmospheric Chemistry and*
797 *Physics*, 7, 815-838, 2007c.



- 798 Seibert, P. and Frank, A.: Source-receptor matrix calculation with a Lagrangian particle dispersion model in
799 backward mode, *Atm. Chem. Phys.*, 4, 51–63, 2004
- 800 Staudt, A. C., Jacob, D. J., Logan, J. A., Bachiochi, D., Krishnamurti, T. N., and Sachse, G. W.: Continental
801 sources, transoceanic transport, and interhemispheric exchange of carbon monoxide over the Pacific, *J. Geophys.*
802 *Res.*, 106(D23), 32571–32590, 2001
- 803 Stein, O., Schultz, M.G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., George, M., Clerbaux, C.: On the
804 wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations, *Atmos.*
805 *Chem. Phys.*, 14, 9295-9316, doi:10.5194/acp-14-9295-2014, 2014
- 806 Stohl, A., M. Hittenberger, and G. Wotawa: Validation of the Lagrangian particle dispersion model FLEXPART
807 against large scale tracer experiments, *Atmos. Environ.*, 32, 4245-4264, 1998
- 808 Stohl, A. and Thomson, D. J.: A density correction for Lagrangian particle dispersion models, *Boundary Layer*
809 *Meteorol.*, 90, 155–167, 1999.
- 810 Stohl, A., Eckhardt, S., Forster, C., James, P., and Spichtinger, N.: On the pathways and timescales of
811 intercontinental air pollution transport, *J. Geophys. Res.*, 107, 2002.
- 812 Stohl, A., Forster, C., Eckhardt, S., Spichtinger, N., Huntrieser, H., Heland, J., Schlager, H., Wilhelm, S.,
813 Arnold, F., and Cooper, O.: A backward modeling study of intercontinental pollution transport using aircraft
814 measurements, *J. Geophys. Res.*, 108, 2003.
- 815 Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion
816 model FLEXPART version 6.2, *Atmos. Chem. Phys.*, 5, 2461-2474, doi:10.5194/acp-5-2461-2005, 2005
- 817 Stohl, A., Aamaas, B., Amann, M., Baker, L. H., Bellouin, N., Berntsen, T. K., Boucher, O., Cherian, R.,
818 Collins, W., Daskalakis, N., Dusinska, M., Eckhardt, S., Fuglestedt, J. S., Harju, M., Heyes, C., Hodnebrog, Ø.,
819 Hao, J., Im, U., Kanakidou, M., Klimont, Z., Kupiainen, K., Law, K. S., Lund, M. T., Maas, R., MacIntosh, C.
820 R., Myhre, G., Myriokefalitakis, S., Ollivier, D., Quaas, J., Quennehen, B., Raut, J.-C., Rumbold, S. T., Samset, B.
821 H., Schulz, M., Seland, Ø., Shine, K. P., Skeie, R. B., Wang, S., Yttri, K. E., and Zhu, T.: Evaluating the climate
822 and air quality impacts of short-lived pollutants, *Atmos. Chem. Phys.*, 15, 10529-10566, doi:10.5194/acp-15-
823 10529-2015, 2015
- 824 Stroppiana, D., Brivio, P. A., Grégoire, J.-M., Lioussé, C., Guillaume, B., Granier, C., Mieville, A., Chin, M.,
825 and Pétron, G.: Comparison of global inventories of CO emissions from biomass burning derived from remotely
826 sensed data, *Atm. Chem. Phys.*, 10, 12173–12189, <http://www.atmos-chem-phys.net/10/12173/2010/>, 2010.
- 827 Tanimoto, H., Ohara, T., Uno, I.: Asian anthropogenic emissions and decadal trends in springtime tropospheric
828 ozone over Japan: 1998-2007, *Geophys. Res. Letters*, doi: 10.1029/2009GL041382, 2009
- 829 Taylor, K. E., Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical*
830 *Research*, 106, D7, 7183-7192, 2001
- 831 Tressol, M., Ordonez, C., Zbinden, R., Brioude, J., Thouret, V., Mari, C., Nedelec, P., Cammas, J.-P., Smit, H.,
832 Patz, H.-W., and Volz-Thomas, A.: Air pollution during the 2003 European heat wave as seen by MOZAIC
833 airliners, *Atm. Chem. Phys.*, 8, 2133–2150, 2008.
- 834 Turquety, S., Logan, J. A., Jacob, D. J., Hudman, R. C., Leung, F. Y., Heald, C. L., Yantosca, R. M., Wu, S.,
835 Emmons, L. K., Edwards, D. P., and Sachse, G. W.: Inventory of boreal fire emissions for North America in
836 2004: Importance of peat burning and pyroconvective injection, *J. Geophys. Res.*, 112, 2007.



- 837 Van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C.,
838 DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation,
839 savanna, forest, agricultural, and peat fires (1997-2009), *Atm. Chem. Phys.*, 10, 11 707–11 735, 2010.
- 840 Tanimoto, H., Ohara, T., Uno, I.: Asian anthropogenic emissions and decadal trends in springtime tropospheric
841 ozone over Japan: 1998-2007, *Geophys. Res. Letters*, 36, L23802, doi:10.1029/2009GL041382, 2009
- 842 Thouret, V., Cho, J.Y.N., Newell, R.E., Larenco, A. and Smit, H.G.J.: General characteristics of tropospheric
843 trace constituent layers observed in the MOZAIC program, *J. of Geophys. Res.*, 105, D13, 17379-17392, doi:
844 10.1029/2000JD900238, 2000
- 845 Thouret, V., Cammas, J.-P., Sauvage, B., Athier, G., Zbinden, R., Nédélec, P., Simon, P., and Karcher, F.:
846 Tropopause referenced ozone climatology and inter-annual variability (1994–2003) from the MOZAIC
847 programme, *Atmos. Chem. Phys.*, 6, 1033-1051, doi:10.5194/acp-6-1033-2006, 2006
- 848 Vay, S. A., , Y. Choi, K. P. Vadrevu, D. R. Blake, S. C. Tyler, A. Wisthaler, A. Hecobian, Y. Kondo, G. S.
849 Diskin, G. W. Sachse, J.-H. Woo, A. J. Weinheimer, J. F. Burkhart, A. Stohl, and P. O. Wennberg : Patterns of
850 CO₂ and radiocarbon across high northern latitudes during International Polar Year 2008. *J. Geophys. Res.* 116,
851 D14301, doi:10.1029/2011JD015643, 2011
- 852 Wang, X., Wang, Y., Hao, J., Kondo, Y., Irwin, M., Munger, J.W., Zhao, Y.: Top-down estimate of China's
853 black carbon emissions using surface observations: sensitivity to observation representativeness and transport
854 model error, *J. of Geophys. Res.*, 118, 5781-5795, doi: 10.1002/jgrd.50397, 2013.
- 855 Wen, D., Lin, J.C., Millet, D.B., et al.: A backward-time stochastic Lagrangian air quality model, *Atmos. Env.*
856 54, 373-386, 2012
- 857 Yamasoe, M.A.; Sauvage, B.; Thouret, V.; et al. : Analysis of tropospheric ozone and carbon monoxide profiles
858 over South America based on MOZAIC/IAGOS database and model simulations, *Tellus B*, 67, 27884, 2015
- 859 Yashiro, H., Sugawara, S., Sudo, K., Aoki, S., and Nakazawa, T.: Temporal and spatial variations of carbon
860 monoxide over the western part of the Pacific Ocean, *J. Geophys. Res.*, 114, D08305, doi:10.1029/2008jd010876,
861 2009
- 862 Zhang, Yiqiang; Liu, Hongyu; Crawford, James H.; et al. : Distribution, variability and sources of tropospheric
863 ozone over south China in spring: Intensive ozonesonde measurements at five locations and modeling analysis, *J.*
864 *of Geophys. Res.*, 117, D12304, 2012
- 865 Zhang, L., Henze, D.K., Grell, G.A. et al.: Constraining black carbon aerosol over Asia using OMI aerosol
866 absorption optical depth and the adjoint of GEOS-Chem, *Atmos. Chem. Phys.*, 15, 10281-10308,
867 doi:10.5194/acp-15-10281-2015, 2015
- 868
869
870
871
872
873
874
875
876



877

878

879

Date	Take-off	Landing	Used for choosing
10 March 2002	Frankfurt	Denver	Anthropogenic emission inventories
27 November 2002	Dallas	Frankfurt	Anthropogenic emission inventories
4 June 2003	Tokyo	Vienna	Fire injection heights (pyro-convection)
6 August 2003	Boston	Frankfurt	Fire injection heights
9 August 2003	Dubai	Frankfurt	Fire injection heights
10 August 2003	Frankfurt	Dallas	Fire injection heights
29 June 2004	Caracas	Frankfurt	Fire injection heights (pyro-convection)
30 June 2004	Frankfurt	Washington	Fire injection heights (pyro-convection) Fire inventories
22 July 2004	Frankfurt	Atlanta	Fire injection heights (pyro-convection) Fire inventories
22 July 2004	Douala	Paris	Fire injection heights (pyro-convection) Fire inventories
23 July 2004	Frankfurt	Atlanta	Fire injection heights (pyro-convection) Fire inventories
19 July 2005	München	Hong Kong	Anthropogenic emission inventories
22 October 2005	München	Hong Kong	Anthropogenic emission inventories
30 July 2008	Windhoek	Frankfurt	Fire injection heights Fire emission inventories
31 July 2008	Frankfurt	Windhoek	Fire injection heights Fire emission inventories

880 **Table 1: Case studies used to define model settings. Cases studies discussed in the manuscript are in bold**

881

882

883

884

885

886

887

888

889

890

891

Inventory	Temporal coverage	Horizontal resolution	Temporal resolution	Reference
	<i>Anthropogenic emissions</i>			
MACCity	1960 – 2014 +	0.5° x 0.5°	Monthly	<i>Granier et al. (2011)</i>



EDGAR v4.2	1970 - 2008	0.5° x 0.5°	Yearly	<i>Janssens-Maenhout et al. (2010)</i>
<i>Biomass Burning emissions</i>				
GFED 4	1997 - 2017+	0.5° x 0.5°	Daily	<i>Giglio et al. (2013)</i>
GFAS v1.0	2002	0.5° x 0.5°	Daily	
GFAS v1.2	2003 - 2017 +	0.1° x 0.1°	Daily	<i>Kaiser et al. (2012)</i>
ICARTT	2004	1° x 1°	Daily	<i>Turquety et al. (2007)</i>

892 Table 2: List of emission inventories used in this study.

893

Flight	MACCity	EDGAR v4.2
10 March 2002 Frankfurt – Denver	+	
27 November 2002 Dallas – Frankfurt	=	=
19 July 2005 München - Hong Kong	+	
22 October 2005 München - Hong Kong	+	

894 Table 3. Summary of optimal inventory (indicated by a plus sign) determined for representing anthropogenic
895 emissions for different case studies. Equal sign indicate that the case is non-conclusive.

896

Flight	GFED 4	GFAS v1.2	ICARTT
29 June 2004 Caracas - Frankfurt			+
30 June 2004 Frankfurt - Washington			+
22 July 2004 Frankfurt - Atlanta			+
22 July 2004 Douala - Paris			+
23 July 2004 Frankfurt - Atlanta			+
30 July 2008 Windhoek - Frankfurt	+		N/A
31 July 2008 Frankfurt - Windhoek	=	=	N/A

897 Table 4. Summary of optimal inventory (indicated by a plus sign) determined for representing fire emissions for
898 different case studies. Equal signs indicate that the case is non-conclusive. Note that the ICARTT inventory is only
899 available for summer 2004.

900

901

902

903

904

905

906

907

908

909

910

911

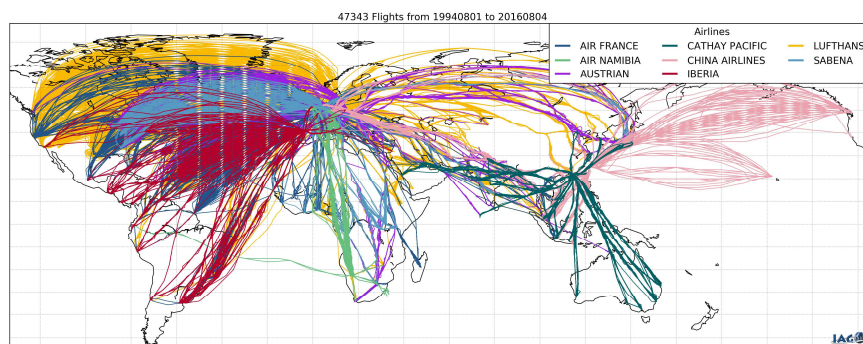
912

913

914

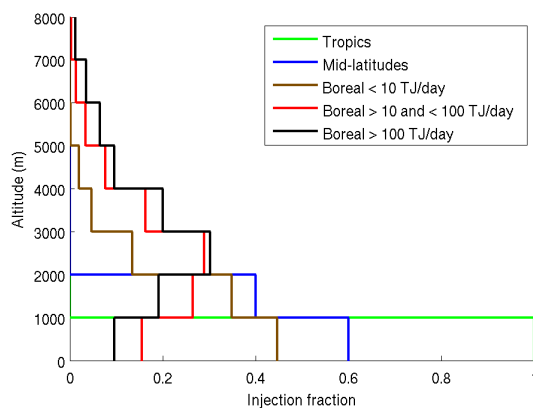
915

916



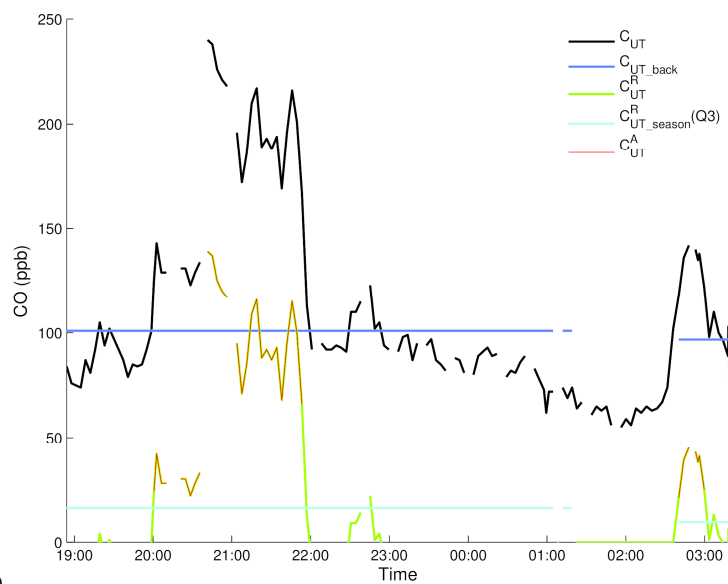
917
918 **Figure 1 : Map showing all flights performed by the IAGOS program**

919
920



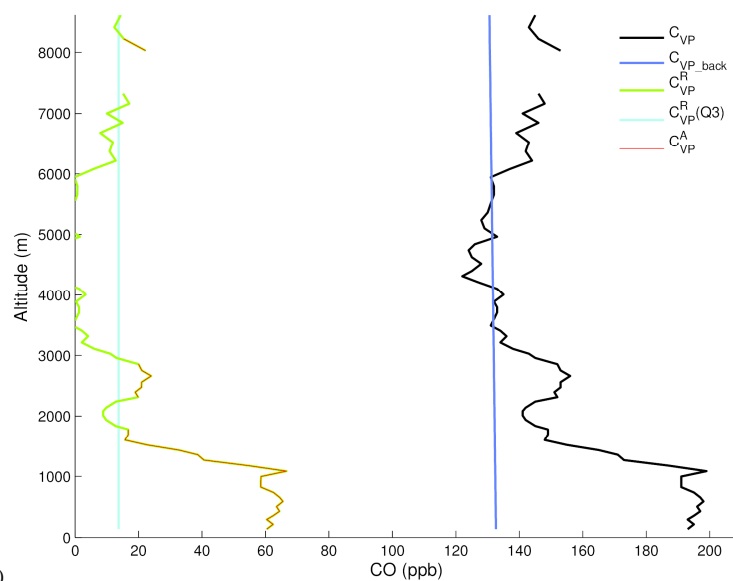
921
922 **Figure 2: Injection profiles used for biomass burning emissions for different regions (Tropics, Mid-latitudes, Boreal)**
923 **in the MIXED methodology.**

924
925



926

a)



927

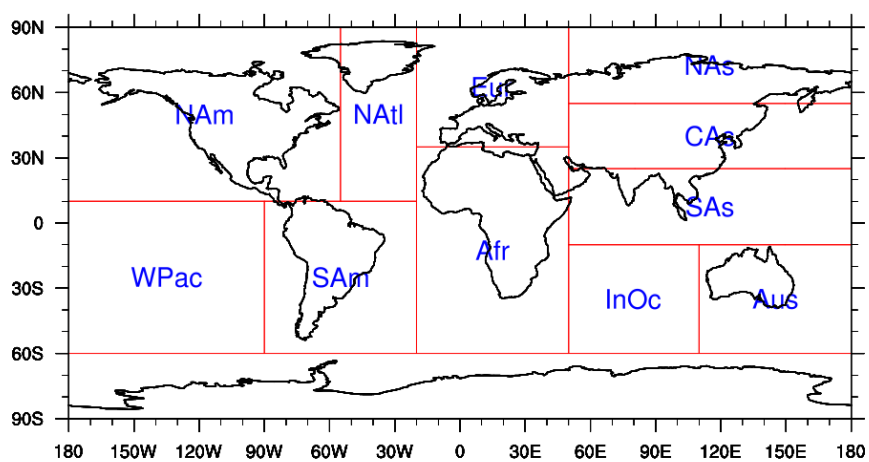
b)

Figure 3: Methodology used to extract CO anomalies along the flight track for (a) the cruise part of the flight and (b) during take off and landing. Further details are given in section 3.4.

928

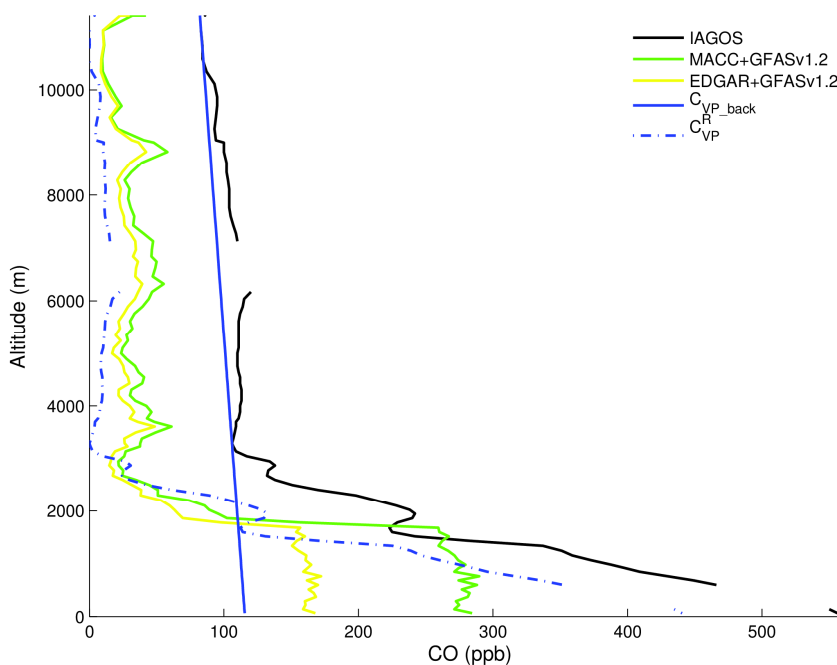
929

930



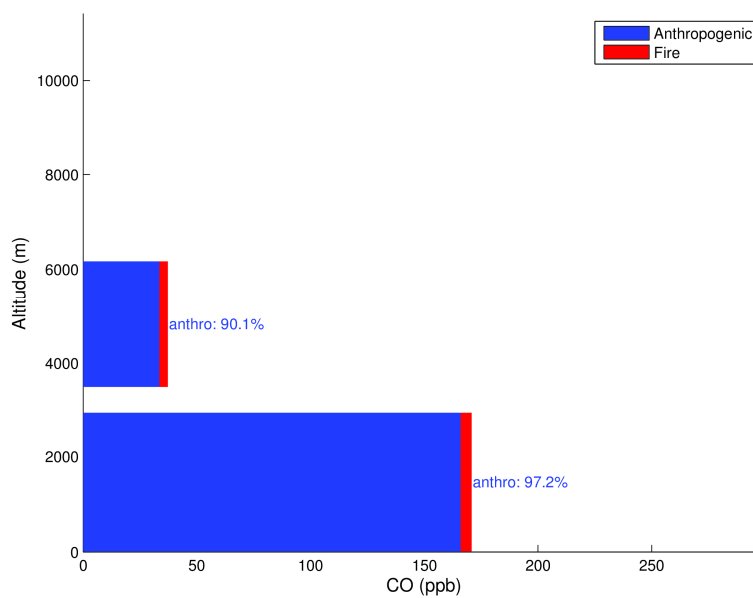
931
932
933
934

Figure 4: Map of the defined regions used to sort IAGOS CO anomalies



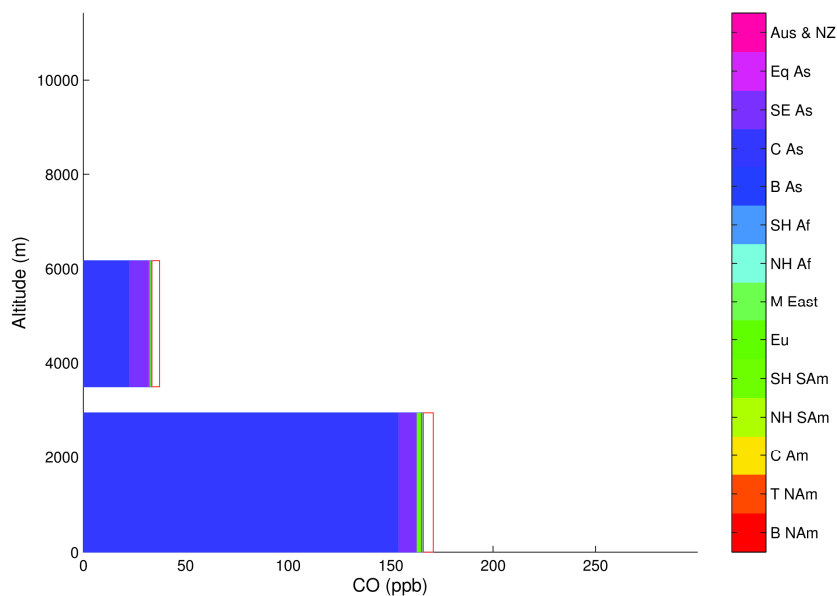
935

a)



936

b)



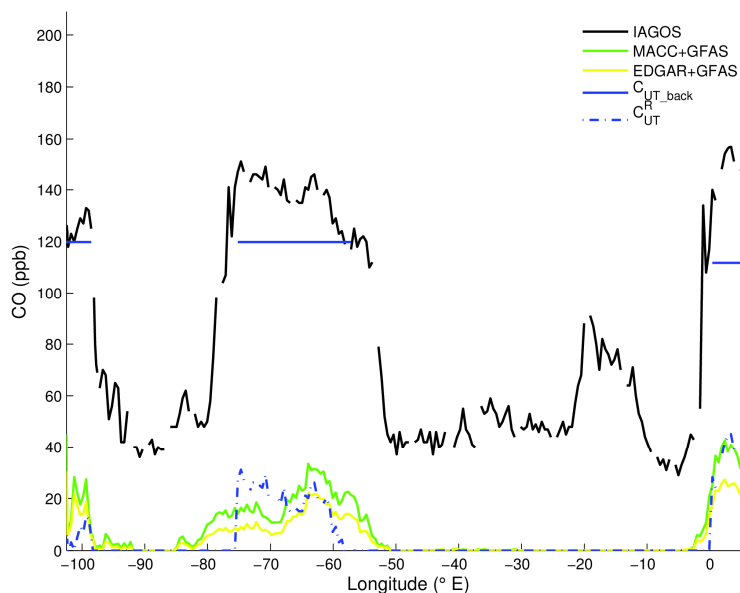
937

c)

938

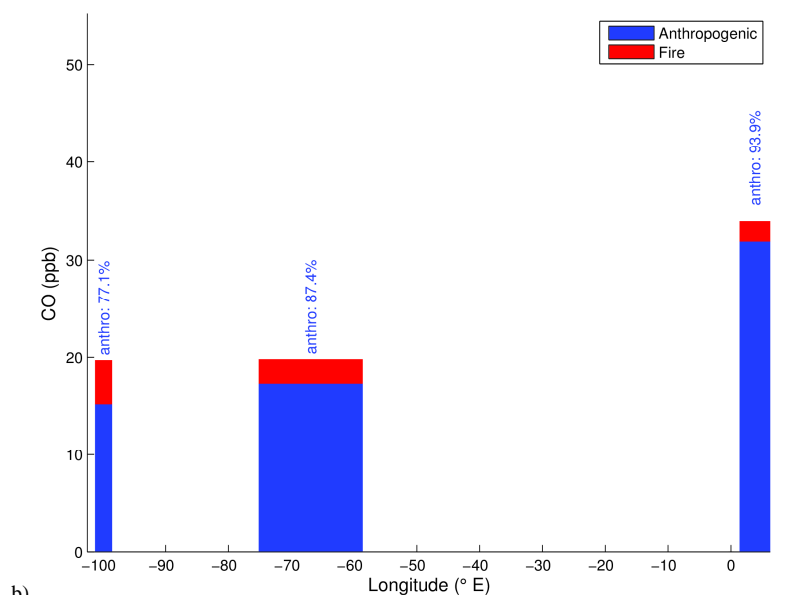
Figure 5: (a) Carbon monoxide profiles over Hong Kong during a MOZAIC-IAGOS flight landing on 22 October 2005. The black line indicates the observed CO profile while the blue line indicates the CO background deduced from the observations. Green and yellow lines indicate the simulated CO contributions using respectively MACCity and EDGARv4.2 for anthropogenic emissions, and using GFAS v1.2 for biomass burning emissions. Simulated CO is separated in (b) sources contribution (anthropogenic in blue, fires in red) and in (c) regional anthropogenic origins (14 regions defined for global emission inventory, <http://www.globalfiredata.org/data.html>; unshaded red square is for fire contribution), using MACCity and GFASv1.2.

945



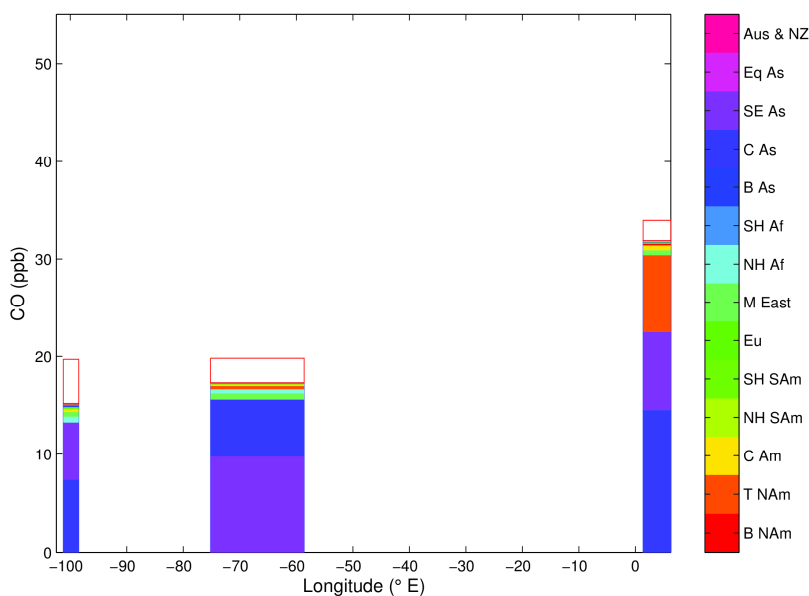
946

a)



947

b)

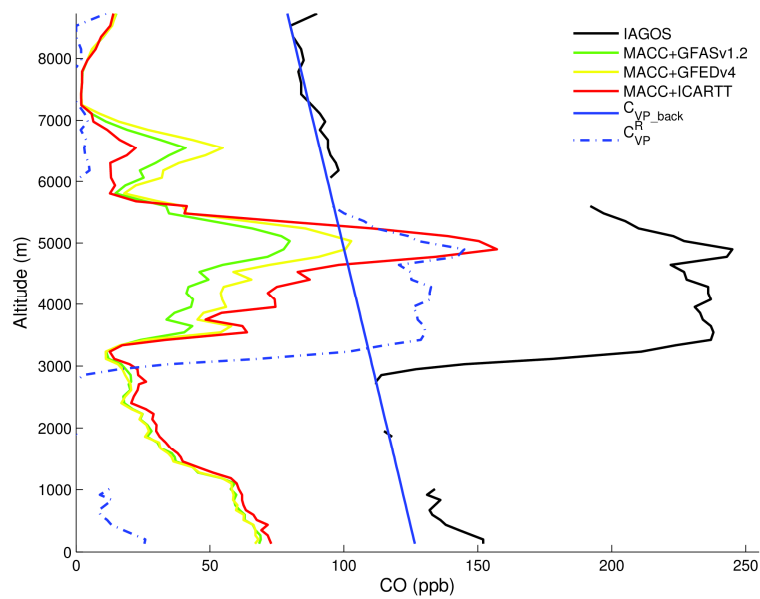


948

c)

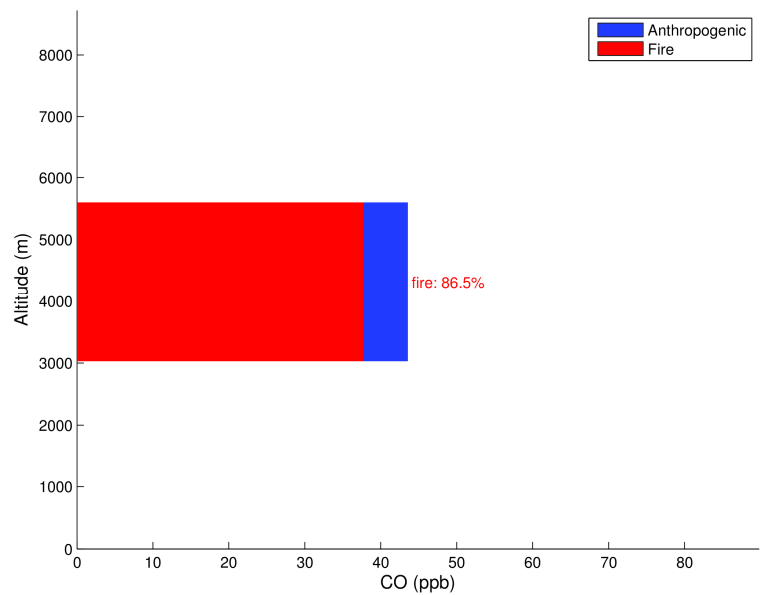
949 **Figure 6:** (a) Carbon monoxide zonal profile during the 10 March 2002 MOZAIC-IAGOS flight from Frankfurt to
 950 Denver. The black line indicates the observed CO while the blue line indicates CO seasonal background in the UT
 951 deduced from the IAGOS data set. Green and yellow lines indicate the simulated contributions using respectively
 952 MACCity and EDGARv4.2 for anthropogenic emissions, and GFAS v1.0 for biomass burning emissions. Simulated
 953 CO is separated in (b) sources contribution (anthropogenic in blue, fires in red) and in (c) regional anthropogenic
 954 origins (14 regions defined for global emission inventory, <http://www.globalfiredata.org/data.html>; unshaded red
 955 square is for fire contribution), using MACCity and GFASv1.0.

956



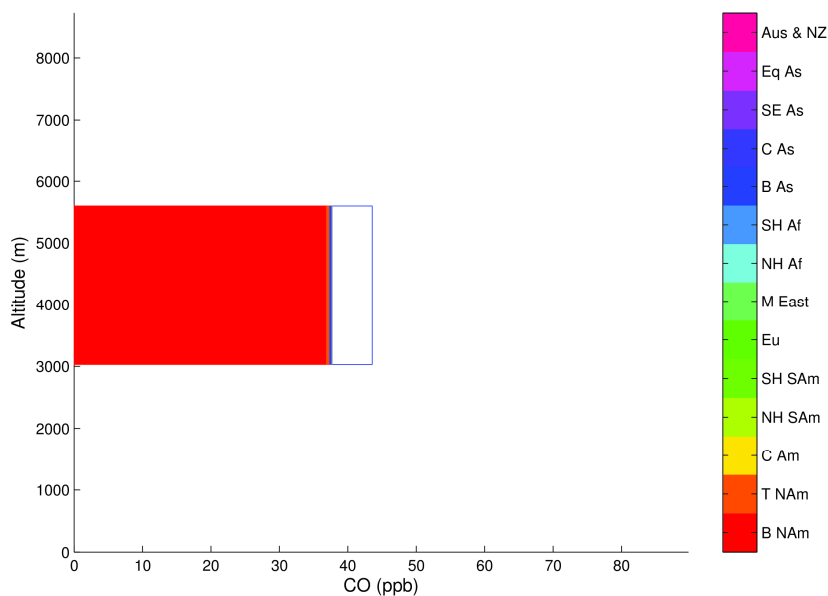
957

a)



958

b)

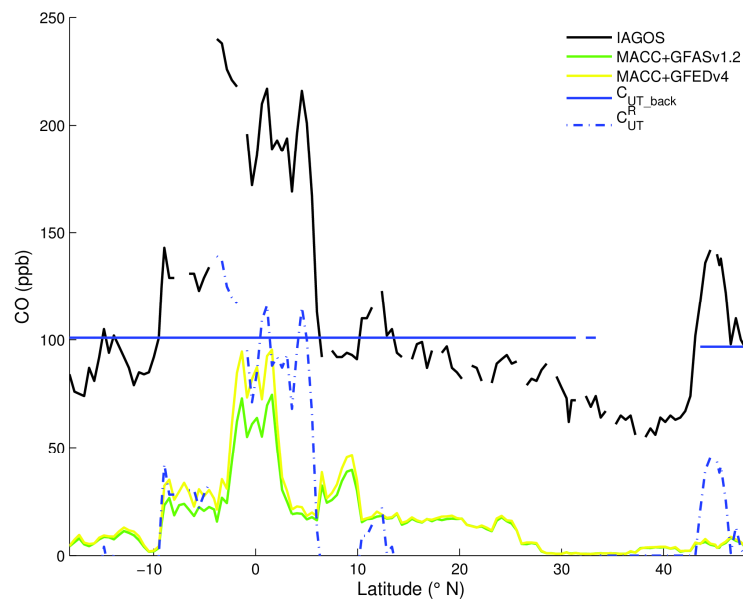


959

c)

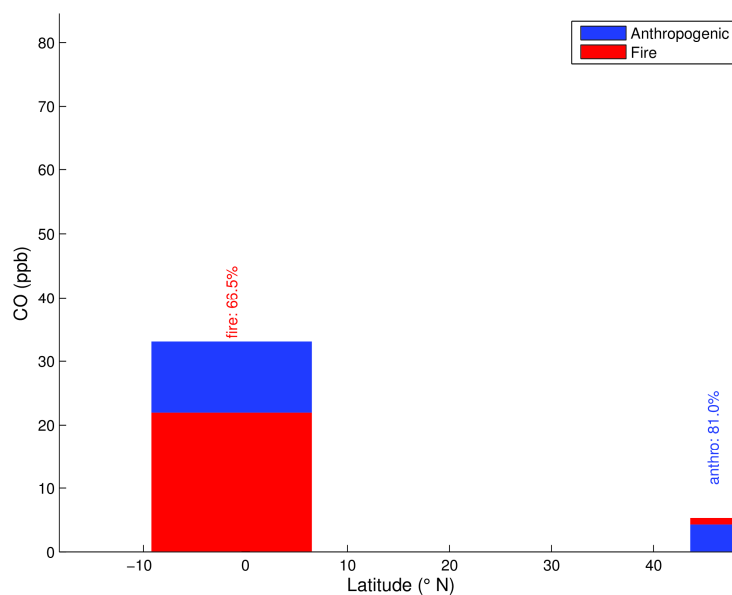
960 **Figure 7 : (a) Carbon monoxide profiles over Paris during a MOZAIC-IAGOS flight landing on 22 July 2004. The**
 961 **black line indicates the observed CO profile and the blue line indicates CO background deduced from the**
 962 **observations. Green, yellow and red lines indicate the simulated contributions using respectively GFASv1.2, GFED4**
 963 **and ICART for biomass burning emissions, with MACCity for anthropogenic emissions. Simulated CO is separated**
 964 **in (b) sources contribution (anthropogenic in blue, fires in red) and in (c) regional biomass burning origins (14 regions**
 965 **defined for global emission inventory, <http://www.globalfiredata.org/data.html>;**
 966 **unshaded blue square is for anthropogenic contribution), using MACCity and GFASv1.2.**

967

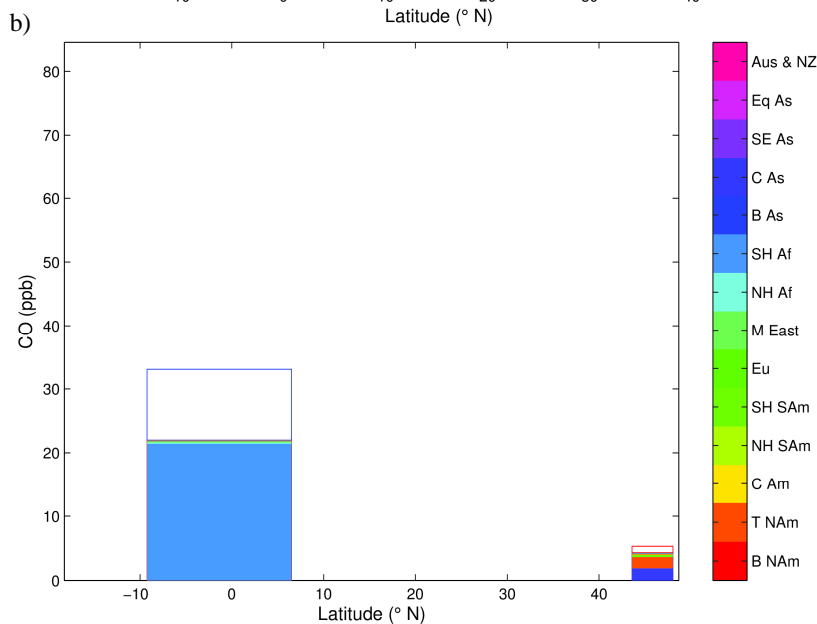


968

a)



969

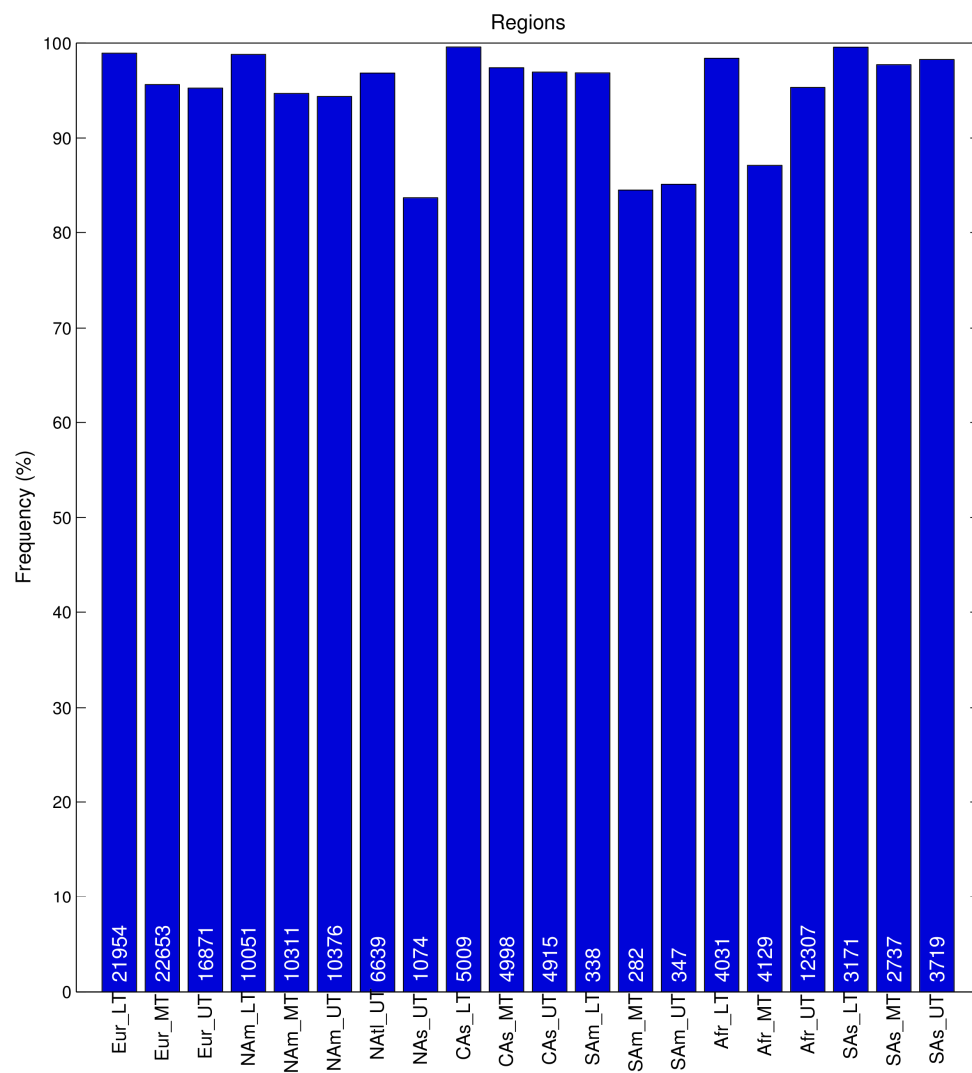


970

c)

Figure 8: (a) Carbon monoxide as a function of latitude during the 30 July 2008 MOZAIC-IAGOS flight from Windhoek to Frankfurt. The black line indicates the observed CO, the blue line indicates the CO seasonal background deduced from the IAGOS data set. Green and yellow lines indicate the simulated contributions using MACCity for anthropogenic emissions, and respectively GFAS v1.2 and GFED4 for biomass burning emissions. Simulated CO is separated in (b) sources contribution (anthropogenic in blue, fires in red) and in (c) regional biomass burning origins (14 regions defined for global emission inventory, <http://www.globalfiredata.org/data.html>; unshaded red square is for anthropogenic contribution), using MACCity and GFASv1.2.

978



979
980

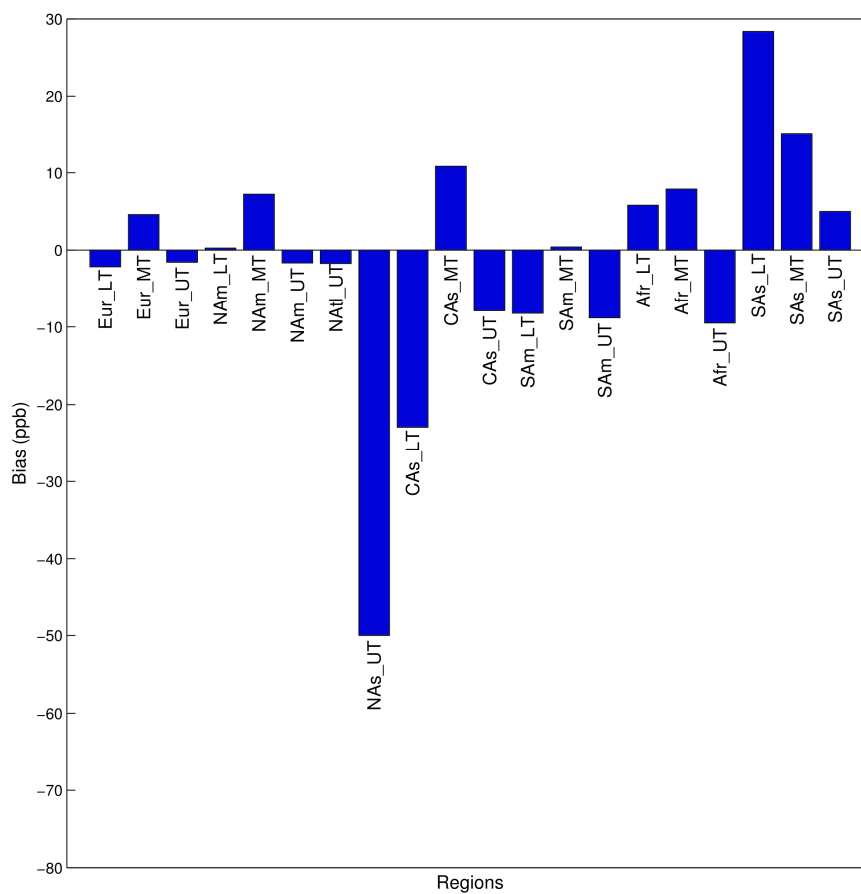
981
982

983 **Figure 9: Frequency of plume detection (a) in different regions / altitudes / seasons using the MACCity and GFAS**
 984 **v1.2 emission inventories during the 2003-2013 period. Biomass burning vertical injection uses APT methodology.**
 985 **Altitude levels stand for LT=0-2km, MT=2-8km and UT=8km-tropopause. The numbers of the plumes observed in**
 986 **each case are displayed in each box.**
 987

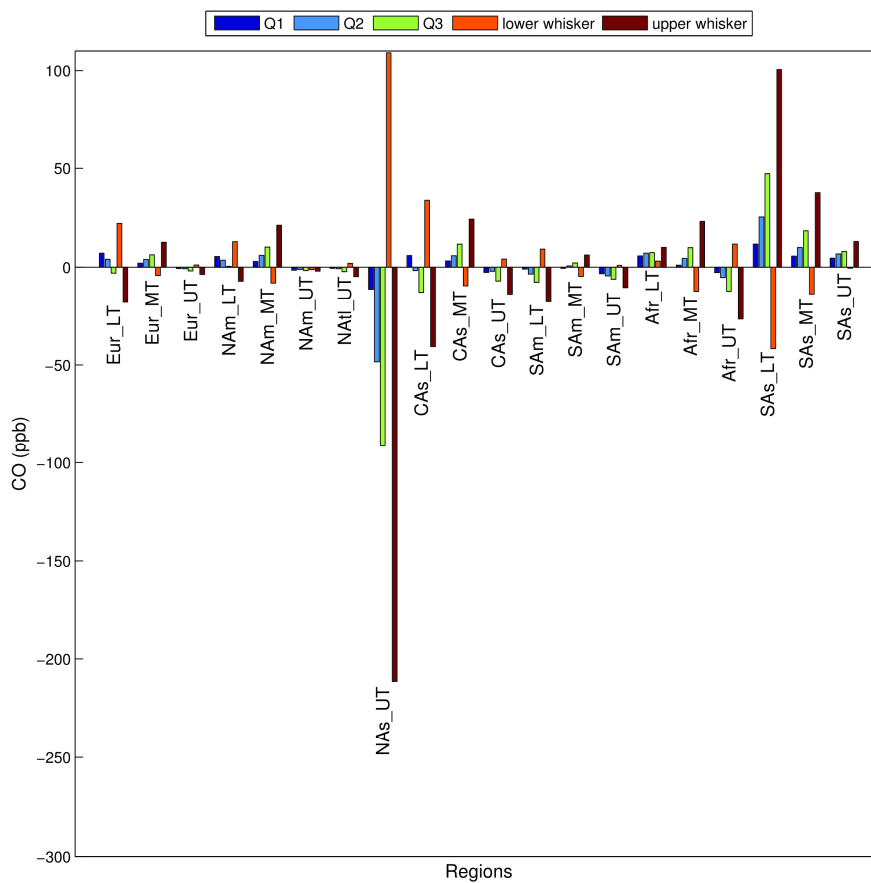
988
989
990
991
992



993
994
995
996
997
998
999
1000
1001
1002
1003
1004

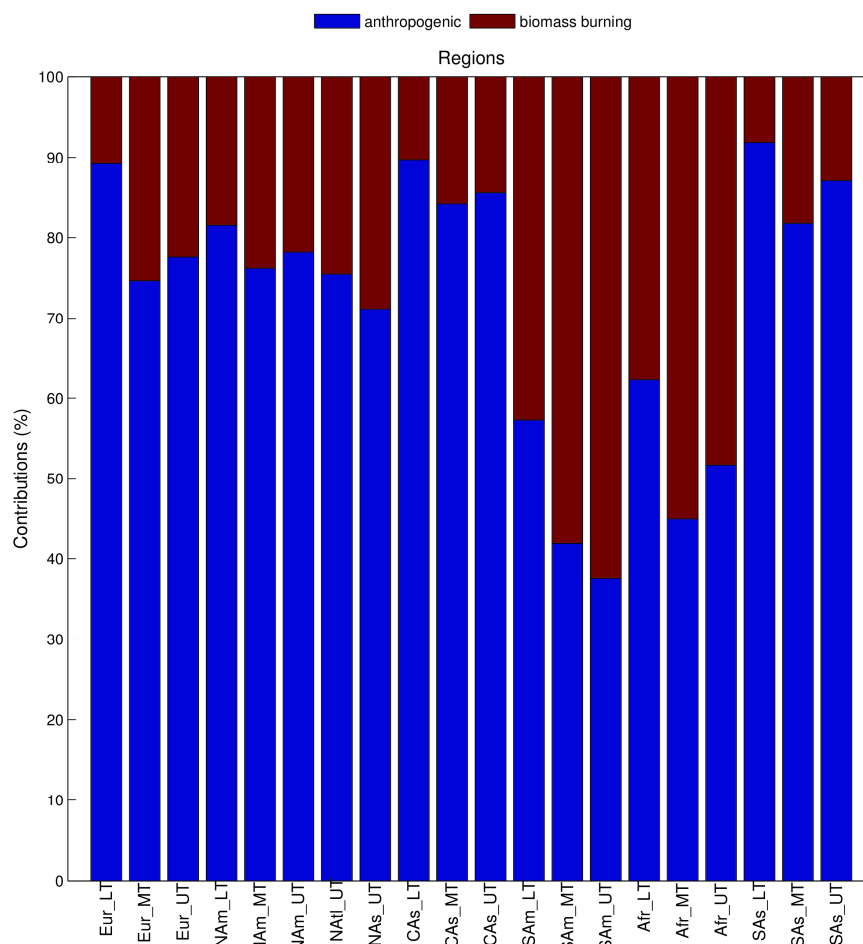


1005



1006

b)

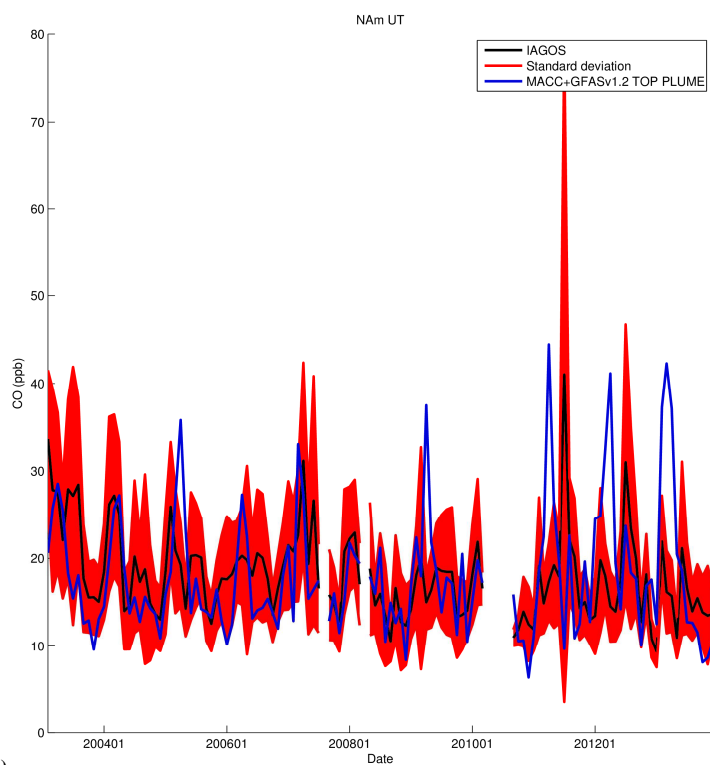


1007

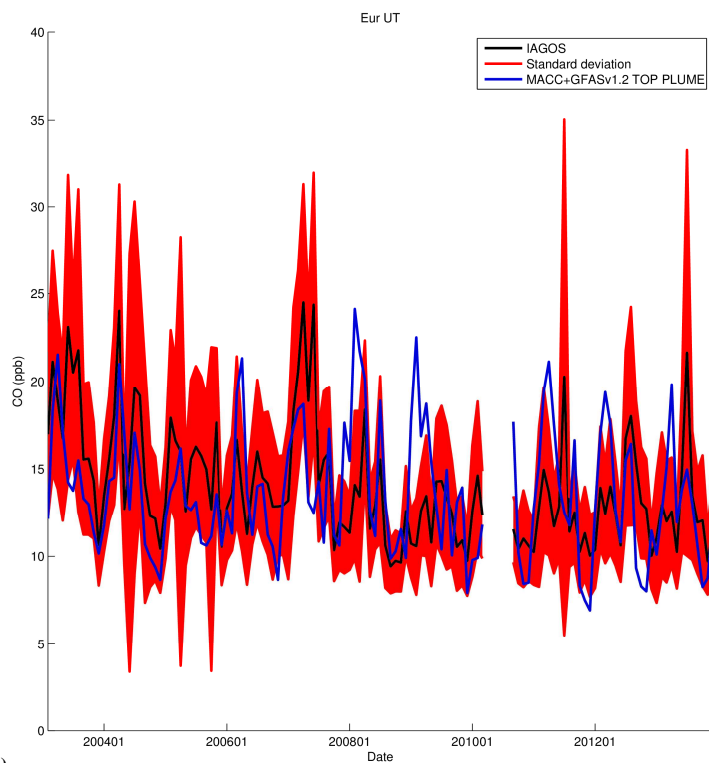
c)

1008 **Figure 10: (a) Mean bias between the modeled and observed CO anomalies ; (b) Percentiles of the modeled CO**
 1009 **anomalies bias with respect to observations; (c) Relative contribution from anthropogenic and biomass burning**
 1010 **sources to the modeled CO. The three graphs are for the main sampled regions (Europe, North America, North**
 1011 **Atlantic, North Asia, Central Asia, South America, Africa, South Asia) and in three layers (LT, MT, UT), using**
 1012 **MACCity and GFASv1.2 for the 2003-2013 period. Biomass burning vertical injection uses APT methodology.**

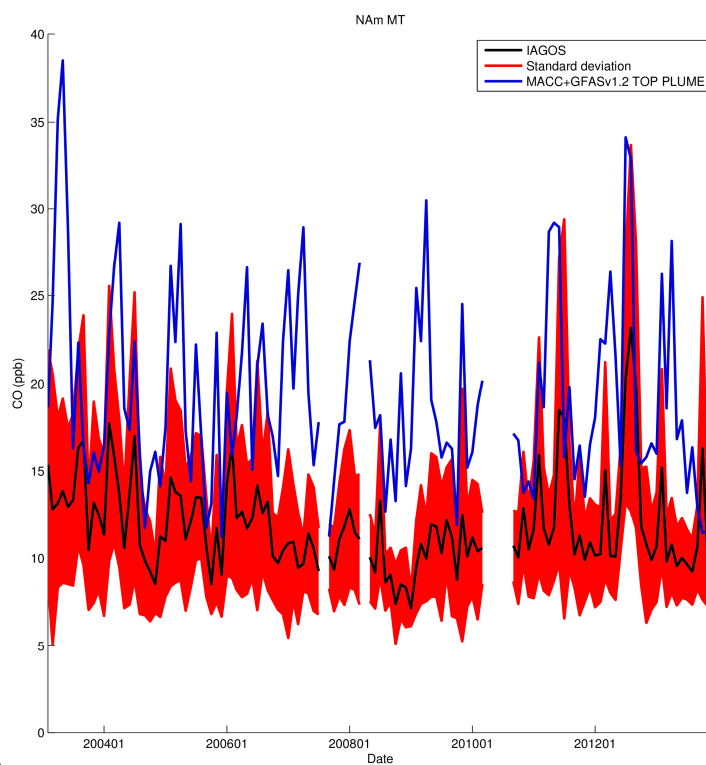
1013



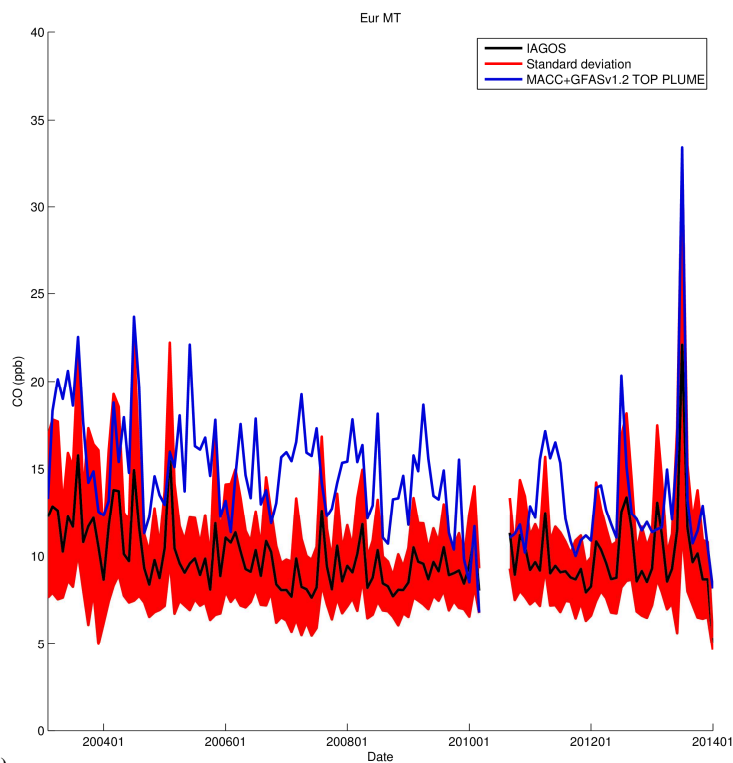
1014 a)



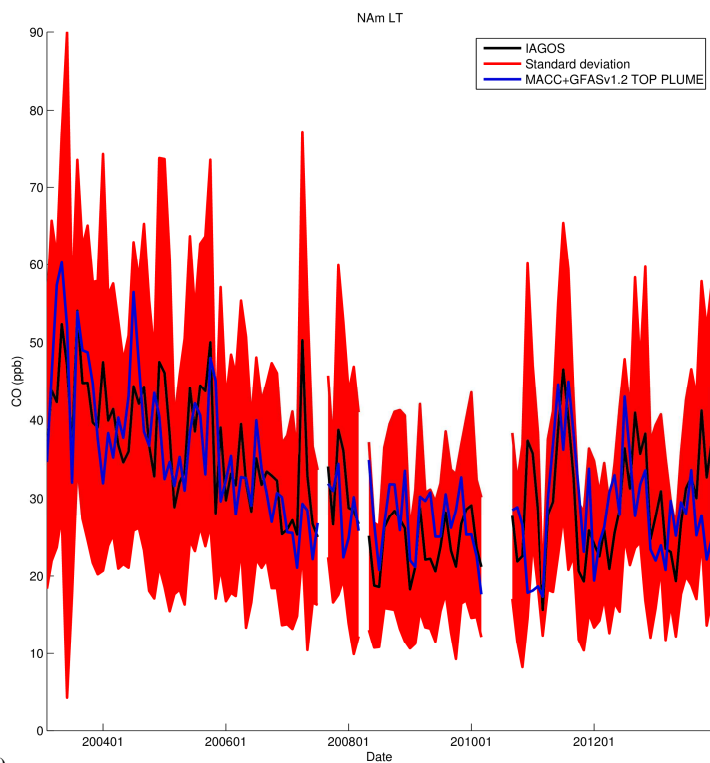
1015 b)



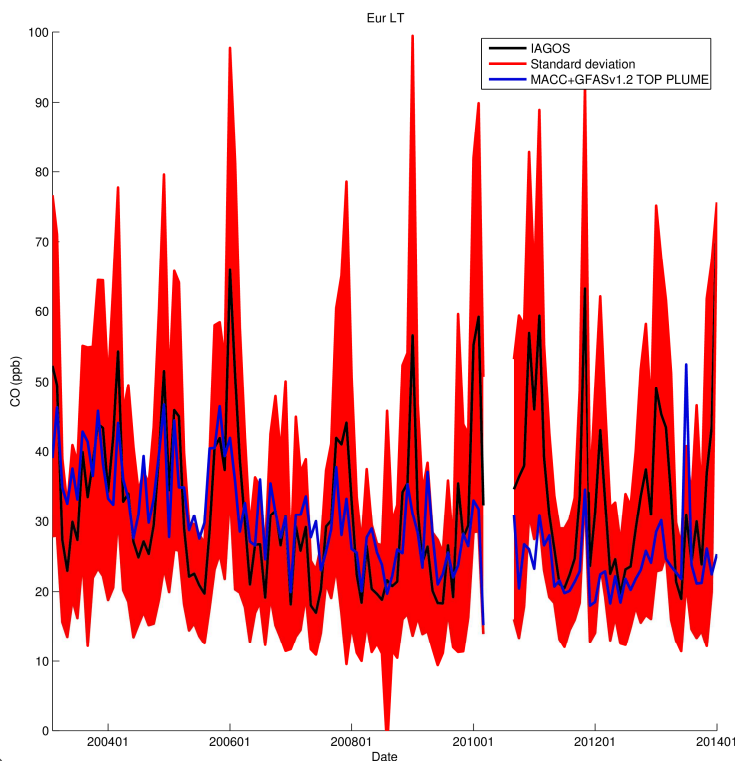
1016 e)



1017 d)



1018 e)



1019

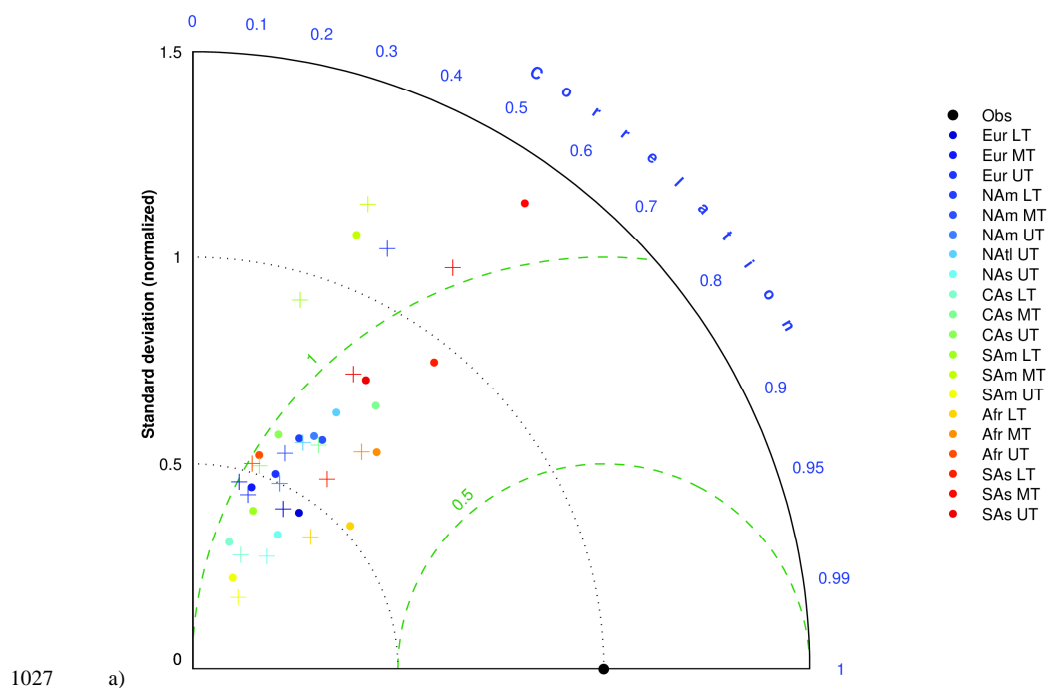
f)

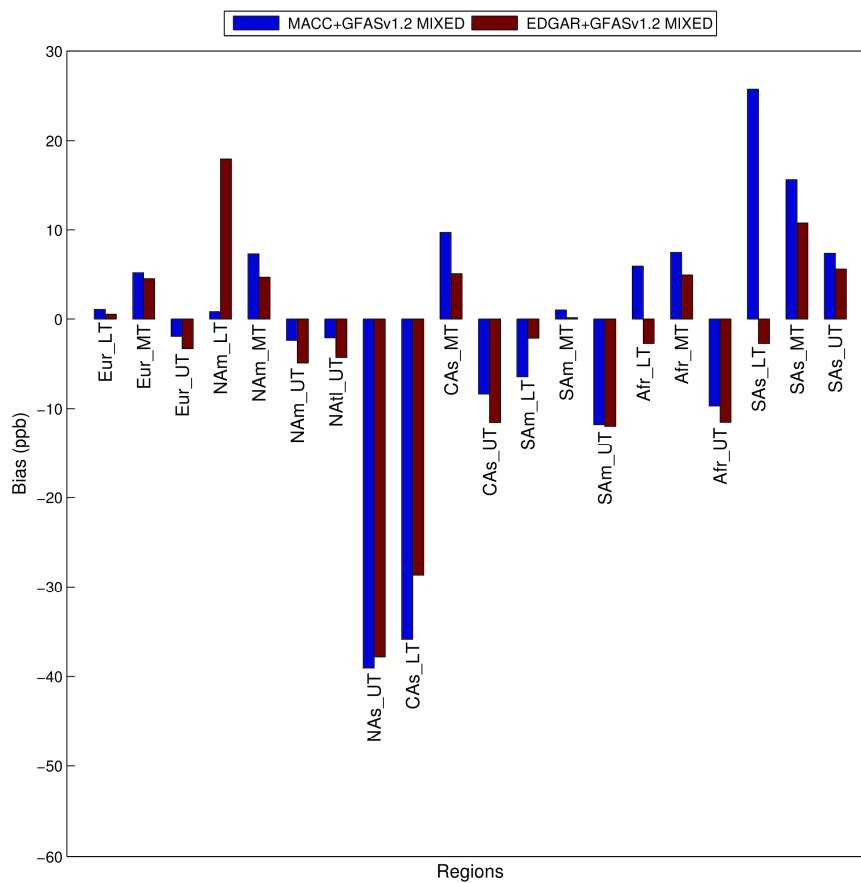
1020 **Figure 11: Times series (monthly means between 2003 and 2013) of the observed (black) and simulated (blue) plumes**
 1021 **of CO enhancements for the two most documented regions (North America and Europe) in the LT (e & f), MT (c & d)**
 1022 **and UT (a & b), using MACCity and GFASv1.2. Biomass burning vertical injection uses APT methodology. Red**
 1023 **shadow represents the standard deviation of the IAGOS observations**

1024

1025

1026





1028

b)

Regions

1029

1030

1031 **Figure 12: Comparison of the SOFT-IO anthropogenic emission influence between 2002 and 2008 (a) Taylor diagrams**

1032 **are obtained for the different regions and in the three vertical layers (LT, MT and UT) using MACCity (dots) and**

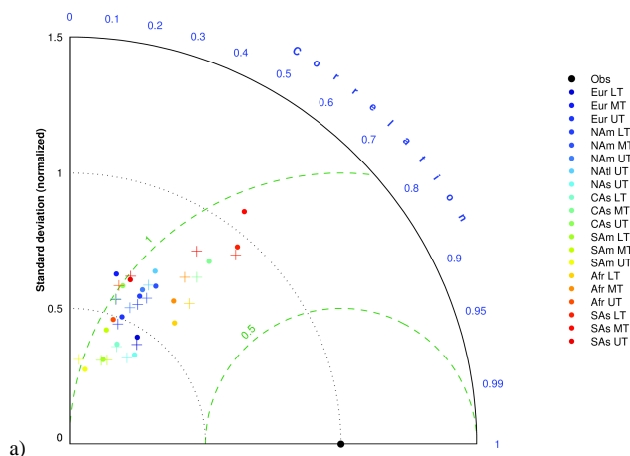
1033 **EDGARv4.2 (crosses) with GFAS (b) Mean biases between the modelled (blue for MACCity + GFAS; brown for**

1034 **EDGARv4.2 + GFAS) and observed CO anomalies. The MIXED methodology is used for fire vertical injection**

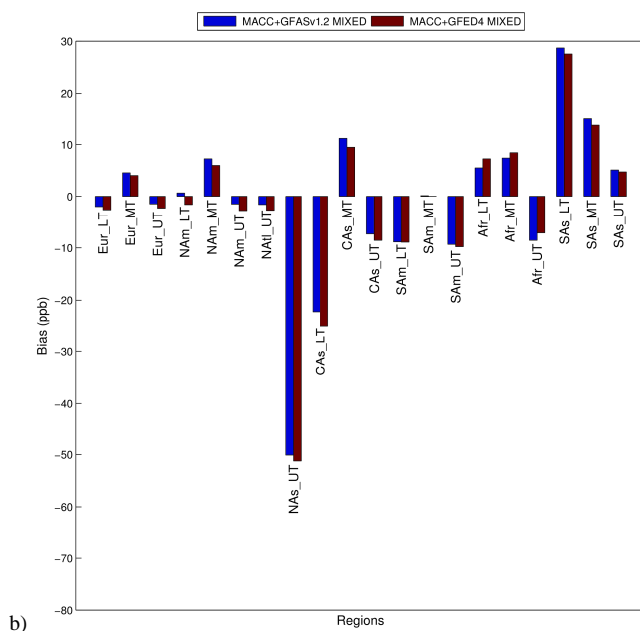
1035



1036

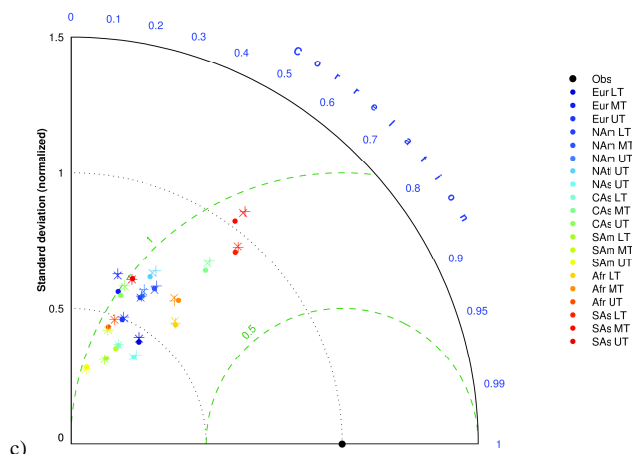


1037



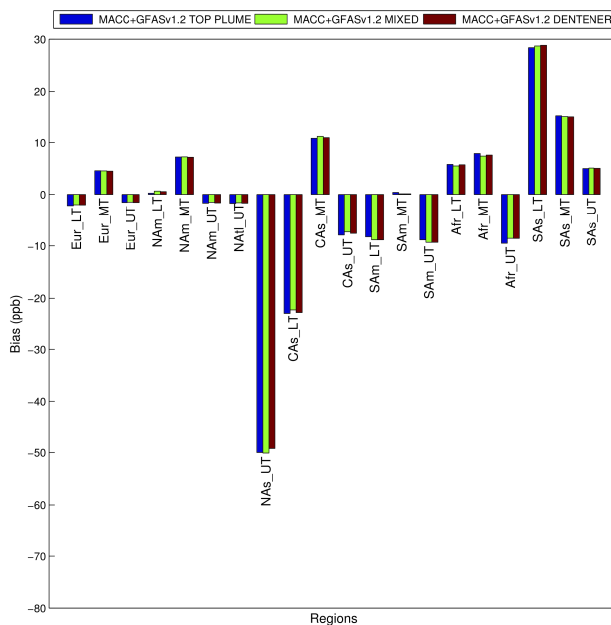


1038



c)

1039



d)

1040

1041 **Figure 13: Comparison of the SOFT-IO biomass burning emission influence between 2003 and 2013. Taylor diagrams**
 1042 **are obtained for the different regions and in the three vertical layers (LT, MT and UT) using (a) GFASv1.2 (dots) and**
 1043 **GFED4 (crosses) with MACCity and MIXED methodology for both GFASv1.2 and GFED4; (c) GFASv1.2 and**
 1044 **MACCity with different vertical fire injections methodologies: MIXED (dots), APT (plus) and DENTENER (crosses).**
 1045 **Mean biases between modeled and observed CO anomalies. Model is using (b) GFASv1.2 + MACCity (blue); GFED4**
 1046 **+ MACCity (brown) and MIXED methodology for both GFASv1.2 and GFED4; (d) GFASv1.2 + MACCity and**
 1047 **different vertical fire injections methodologies: MIXED (blue); APT (green) and DENTENER (brown)**

1048

1049

1050

1051